Acquisition Directorate
Research & Development Center

Report No. CG-D-09-16

Deterrence Impact Modeling Environment (DIME) Proof-of-Concept Test Evaluations and Findings

June 2016

Homeland Security

N O T I C E

This document is disseminated under the sponsorship of the Department of Homeland Security in the interest of information exchange. The United States Government assumes no liability for its contents or use thereof.

The United States Government does not endorse products or manufacturers. Trade or manufacturers' names appear herein solely because they are considered essential to the object of this report.

This report does not constitute a standard, specification, or regulation.

Mr. Timothy Girton
Technical Director
United States Coast Guard
Research & Development Center
1 Chelsea Street
New London, CT 06320

Technical Report Documentation Page

1. Report No. CG-D-09-16	**2. Government Accession Number**	**3. Recipient's Catalog No.**
4. Title and Subtitle Deterrence Impact Modeling Environment (DIME) Proof-of-Concept Test Evaluations and Findings		**5. Report Date** June 2016 **6. Performing Organization Code** Project No. 7525
7. Author(s) Mr. Phillip J. Palin, Dr. Rodrigo Nieto-Gomez, Dr. Jamison Day, Mr. Craig W. Baldwin		**8. Performing Report No.** R&DC UDI # 1402
9. Performing Organization Name and Address Center for Homeland Defense and Security (CHDS) Naval Postgraduate School 1 University Circle Bldg 220, Rm. 064 Monterey, CA 93943	U.S. Coast Guard Research and Development Center 1 Chelsea Street New London, CT 06320	**10. Work Unit No. (TRAIS)** **11. Contract or Grant No.** MIPR# HSCG32-13-X-R00011
12. Sponsoring Organization Name and Address COMMANDANT (CG-DCO-81) US COAST GUARD STOP 7318 2703 MARTIN LUTHER KING JR AVE SE WASHINGTON, DC 20593		**13. Type of Report & Period Covered** Final Report **14. Sponsoring Agency Code** Commandant (CG-DCO-81) US Coast Guard Stop 7318 Washington, DC 20593

15. Supplementary Notes

The R&D Center's technical point of contact is Craig W. Baldwin , 860-271-2652, email: Craig.W.Baldwin@uscg.mil

16. Abstract (MAXIMUM 200 WORDS)

The exploding use of social media and digital monitoring presents the Coast Guard with a crucial new domain for mission operations. At the very least, the digital domain is a source of situational awareness for maritime operations. Over time and with creative engagement, the digital domain offers the Coast Guard potentially powerful tools to intelligently inform policy, strategy and planning decision making across most – even all – mission sets.

The study explores the Data Driven Decision Making Cycle metaphorical concepts of a Digital Ocean and a pilot project as means to contextually define the exploding data streams associated with the emergence of the socio-technological domain and a means for engaging this domain. The study shows how a pilot project institutionalizes a Coast Guard capability to build tools (mobile applications) that: identify data escapes where digital pheromones are being produced; capture them; identify trends and patterns; and produce a mechanism that allows decision makers to visualize and decide where, when and how to intervene, as well as visualize the results of that intervention.

The Data Driven Decision-Making cycle is developed to provide the pilot project with an information management framework that supports decision-making at all levels by:

- Identifying emerging patterns of strategic importance.
- Routing relevant information among tactical decision support tools.
- Supporting continuous improvement of operational capabilities.

17. Key Words Coast Guard, cyber, deterrence, big-data decision-making.	**18. Distribution Statement** Distribution Statement A: Approved for public release; distribution is unlimited.		
19. Security Class (This Report) UNCLAS//Public	**20. Security Class (This Page)** UNCLAS//Public	**21. No of Pages** 60	**22. Price**

(This page intentionally left blank.)

EXECUTIVE SUMMARY

A 2012 study of Coast Guard policy and practice by a team from the Research and Development Center and the Naval Postgraduate School outlined an effective Coast Guard culture of deterrence. The study also suggested that recent socio-technological advances offered the potential to translate this culture into an explicit strategy and process. During 2014-2015, the team examined the near-term opportunity to implement such a strategy and process.

The exploding use of social media and digital monitoring presents the Coast Guard with a crucial new domain for mission operations. At the very least, the digital domain is a source of situational awareness for maritime operations. Over time and with creative engagement, the digital domain offers the Coast Guard potentially powerful tools to intelligently inform policy, strategy, planning, and decision making across most (if not all) mission sets. To date, however, the Coast Guard has been cautious engaging the new domain. There are several sources of this caution: financial, technical, legal, and ethical.

Several current Coast Guard policies complicate ongoing engagement with and assessment of the digital domain. Policies designed to secure Coast Guard communications and to protect the individual privacy of stakeholder communities, among other goals, have had the unintended consequence of discouraging creative exploration of the digital domain. Security and privacy are each crucial issues, but so is meaningful engagement with what might be characterized as a strategic third area of command in the Coast Guard: Atlantic Area, Pacific Area, and the allegorical Digital Area. Indeed, the Coast Guard's Cyber Strategy[1] has defined "enabling operations" as one of three strategic focus areas.

There is substantive evidence that by deploying Big Data analytics, the Coast Guard could enhance several aspects of mission operations, especially deterrence. The Coast Guard is not, however, currently poised to move in this direction. Today the digital domain is viewed mostly as an external affairs activity. The results of this study suggest the digital domain is, instead, an essential context and potential tool for achieving safety, stewardship and security missions. With the new Cyber Strategy in place, the Coast Guard has taken the first step to change this paradigm.

This report describes an environment where the digital space is fully engaged at the strategic, operational and tactical levels. The Deterrence Integration Modeling Environment is powered by the Data Driven Deterrence (D^3) Cycle.

The D^3 Cycle describes a deliberative process that facilitates analysis of influences that Coast Guard multidimensional presence can have on Social Identity Groups of interest, by tracking idea creation, group adoption and their manifestation in beliefs, behaviors and attitudes expressed not only in the tactical and operational maritime environments but through the burgeoning socio-technological environments.

[1] *United States Coast Guard Cyber Strategy*, United Stated Coast Guard, June 2015

(This page intentionally left blank.)

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF ACRONYMS AND ABBREVIATIONS

AES	Advanced Encryption Standard
AIS	Automatic Identification System
API	Application Program Interface
CAC	Common Access Card
CEO	Chief Executive Officer
COCA	Corpus of Contemporary American English
CPR	Common Pool Resources
CRADA	Cooperative Research and Development Agreement
D^3	Data Driven Decision-making/Data Driven Deterrence
DCS	Data Centric Security
DHS S&T	Department of Homeland Security Science and Technology
DIME	Deterrence Integration Modeling Environment
DoD	Department of Defense
DSC	Digital Selective Calling
FAA	Federal Aviation Administration
FIPS	Federal Information Processing Standard
HSPH	Harvard School of Public Health
IBM	International Business Machines
IoT	Internet of Things
ISN	Instrumental, Social and Normative (a framework)
MPCR	Marginal per Capita Return
NAIS	National Automated Identification System
NPS/CHDS	Naval Postgraduate School/Center for Homeland Defense and Security
PIV	Personal Identity Verification
RDC	Research and Development Center
ROI	Return on Influence
SIG	Social Identity Groups
SIP	Statistically Improbable Phrases
SMM	Social Media Monitoring
USCG	United States Coast Guard
VHF	Very High Frequency
VP	Vice President

(This page intentionally left blank.)

1 INTRODUCTION

1.1 Background and Overview

A 2012 study of Coast Guard policy and practice by a team from the Research and Development Center (RDC) and the Naval Postgraduate School (NPS) outlined an effective Coast Guard culture of deterrence. The study also suggested recent socio-technological advances offered the potential to translate this culture into an explicit strategy and process. During 2014-2015, the team examined the near-term opportunity to implement such a strategy and process.

What the team has been able to show is the ability to perform word scrapes (word counts) on a collection of commercial fishing social media sources and apply some very simplistic computational linguistic heuristics to identify Statistically Improbable Phrases (SIPs) and through association, identify Social Identity Groups (SIGs). A SIG is a population that is characterized by certain linguistic, social, and normative attributes. Knowing these subcomponents, the Coast Guard can develop a deeper understanding of, and may have an opportunity to more effectively engage the SIG.

If this capability is matured, the Coast Guard will find that SIG's will begin to emerge, even within what may appear to be a monolithic SIG. The Coast Guard can apply additional computational behavioral models to measure the level of alignment to beliefs, attitudes and behaviors (good or bad). Over an extended collection timeframe, the team believes the system will be capable of measuring perturbations in alignment and associate them with specific events, actions, regulations and enforcement actions for example. If this is so, the team believes the Coast Guard will be able to define and measure Return on Influence (ROI) from Coast Guard interactions with populations of interest.

1.2 What is the Digital Ocean?

A digital ocean is emerging. It began to form at least three decades ago. Over the last decade it has grown rapidly. In the last five years an unprecedented flood of data has transformed our world. The transformation will continue and accelerate.

This new digital ocean is immediately adjacent to the physical oceans in which the Coast Guard has long operated. The new ocean increasingly reflects and influences how humans behave on the ancient oceans. The new ocean – featuring powerful currents of social media, industrial data, and socio-technical relationships – will enhance both risk and reward in the maritime domain.

The new ocean has strategic implications for safety, stewardship, and security. Coast Guard stakeholders and adversaries are already active there. They will become more active. One day before the results of this study were briefed at Coast Guard headquarters, a new Cyber Strategy was released. It states, "…the Coast Guard will fully embrace cyberspace as an operating domain." Cyberspace is another way of describing our new ocean. As we try to set out in the following, the fundamental character of the new ocean is socio-technical, relational, and profoundly human.

During the past 20+ years the digital ocean has been expanding at an accelerated rate. The last 10 years have greatly accelerated the rate at which data is being generated. In today's social media-driven world, trillions of digital artifacts are being generated and stored each day. Experts in the field forecast that by the year 2020 the annual data creation rate will exceed 45 zeta bytes[2].

To a large extent, the socialization of the Internet is responsible for driving this rate of data creation. This process of Internet socialization was well underway by the late 1990s, but in the last five years there has been explosive growth. The first Apple iPhone was released in 2007. Android phones became available in 2008, but did not achieve significant market penetration until 2010. In 2014 over 1.3 billion new smartphones were shipped worldwide. According to the Pew Research Center in 2014:

- 90% of American adults had a cell phone.
- 58% of American adults had a smartphone.
- 32% of American adults had an e-reader.
- 42% of American adults had a tablet computer.

The Internet of things has evolved and by 2020 is forecasted to consist of billions of devices connected to the Internet. A majority of the devices will consist of sensors, actuators, and controllers while the minority of devices will be made up of computers, tablets and smartphones.

The Internet of Things – plus high-powered analytics – will support much more accurate tracking, prediction, and early intervention in the engineered environment. The same analytic tools can generate similar benefits where digital socialization meets the Internet-of-Things.

Human behavior has always been susceptible to influence. The most powerful sources of influence have usually been the most intimate: family, friends, neighbors, and co-workers. Most humans are predisposed to seek social approval and avoid social disapproval. The emerging digital environment is transforming sources of intimacy and social influence which the Coast Guard can utilize to improve mission effectiveness and efficiency. (Note: A more extensive discussion of what the Digital Ocean is and why it is relevant to this study and Coast Guard operations can be found in Appendix A.)

[2] Source Oracle 2012

2 THE DATA DRIVEN DECISION CYCLE

At its core, the study shows how a Coast Guard capability to build tools (mobile applications) could: identify data escapes where digital pheromones are being produced; capture them; identify trends and patterns; and produce a mechanism that allows decision makers to visualize and decide where, when and how to intervene, as well as visualize the results of that intervention.

Applications can be created to solve one or multiple problems on the Experimental Data Driven Decision Making Cycle (See Figure 1 below). Data apps drive the experimental nature of the cycle.

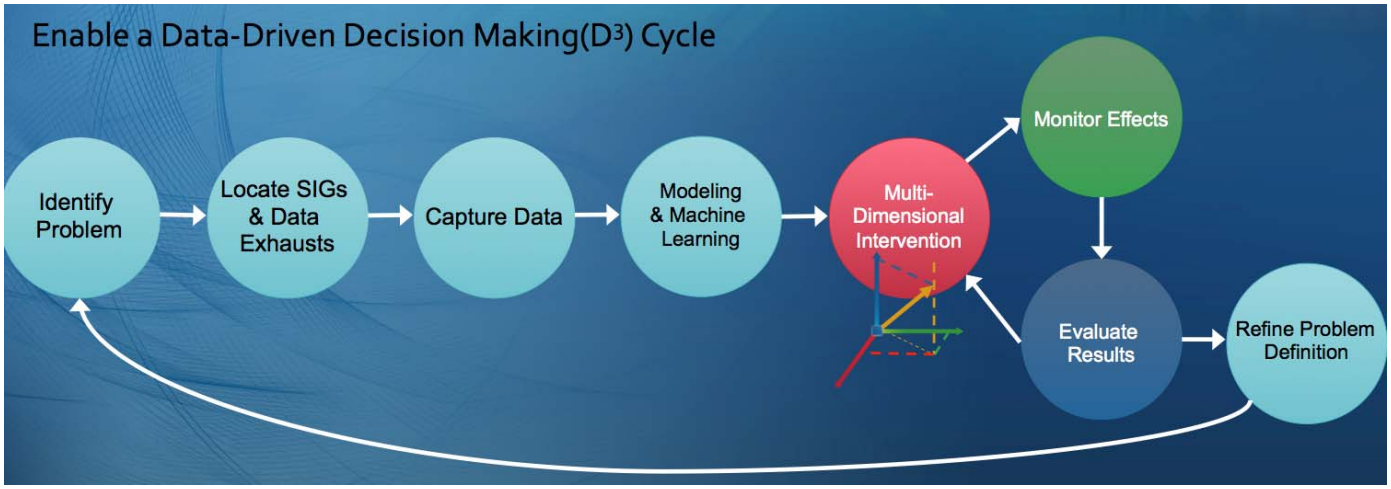


Figure 1. Experimental Data Driven Decision (D^3) making cycle.

As maritime communities are also increasingly digital communities; included in these are Social Identity Groups (SIGs). Within communities, instrumental, social, and normative sanctions continue to be expressed and experienced, but often in ways totally unknown and unavailable to the Coast Guard. Application of Big Data with analytics will allow the Coast Guard to better understand and participate constructively in the digital domain. *Data Driven Decision-Making* or more narrowly, *Data Driven Deterrence* requires **processes** that not only collect, share, and analyze, but also fuse structured data with unstructured information to improve the effectiveness of deterrence efforts.

KEY POINT

The Data Driven Decisions/Deterrence (D^3) cycle is about developing these requisite processes, *not* any specific technologies. Ultimately, D^3 provides an information management framework that supports decision-making at all levels by:

- Identifying emerging patterns of strategic importance.
- Routing relevant information among tactical decision support tools.
- Supporting continuous improvement of operational capabilities.

Several D^3 components were explored and advanced during the research regarding the Deterrence Integration Modeling Environment (DIME) that was conducted in 2014-2015. These components include:

1) **Geo-Temporal Visualization Dashboard** – Enables interactive visualization of data containing location and time elements, promoting rapid identification of space-time patterns.
2) **Mobile Data Collection Tool** – Quickly integrates with existing data management systems to enable high-quality structured data collection via mobile devices.
3) **Data-Centric Security Framework** – Provides auditable assurance of security suitable for allowing government transmission of controlled information via commercial mobile networks and public WiFi. See Appendix C for more detailed information
4) **SIG Identification via Social Media** – Analyzes public and open-source unstructured information to identify USCG deterrence-related SIGs and monitor changes in their behaviors/interests. See Appendix B for more detailed information.

KEY FINDING

Together, these components provide a data-oriented proof of concept for how D^3 processes can use both structured and unstructured information to inform Coast Guard strategic and operational mission planning and execution.

3 AN ANALTICAL REPORT OF THIS INVESTIGATION

The exploding use of social media and digital monitoring presents the Coast Guard with a crucial new domain for mission operations. At the very least, the digital domain is a source of situational awareness for maritime operations. Over time and with creative engagement, the digital domain offers the US Coast Guard potentially powerful tools to intelligently inform policy, strategy and planning decision making across most – even all – mission sets. To date, however, the Coast Guard has been cautious engaging the new domain. There are several sources of caution: financial, technical, legal, and ethical.

There is evidence that several of the stakeholder communities most important to the Coast Guard have not been early adopters of the digital tools that are creating this new domain. But as technology advances and a new generation of stakeholders arise, this is likely to change. As a result, while the Coast Guard is not irretrievably behind-the-curve in terms of the digital domain, there is the risk of a dangerous "digital gap" widening between the current capabilities of the Coast Guard and those who use – and may abuse – the maritime domain.

Several current Coast Guard policies complicate ongoing engagement with and assessment of the digital domain. Policies designed to secure Coast Guard communications and to protect the individual privacy of stakeholder communities, among other goals, have had the unintended consequence of discouraging creative exploration of the digital domain. Security and privacy are each crucial issues, but so is meaningful engagement with what might be characterized as a strategic third area: Atlantic Area, Pacific Area, and Digital Area. Indeed, the Coast Guard's new Cyber Strategy has defined "enabling operations" as one of three strategic focus areas.

There is substantive evidence that by deploying Big Data analytics the Coast Guard could enhance several aspects of mission operations, especially deterrence. The Coast Guard is not currently poised to move in this direction. Today the digital domain is viewed mostly as an external affairs activity. The results of this study suggest the digital domain is, instead, an essential context and potential tool for achieving safety, stewardship, and security missions. With the new Cyber Strategy in place, the Coast Guard has taken the first step to change this paradigm.

There is an urgent need for the Coast Guard to embark on a sustained process of exploring: new policies; new strategies; new planning and training activities; and new tactics, techniques and procedures that effectively engage the digital domain in ways that are clearly: constitutional; consistent with the values of maritime communities; technically and financially sustainable; and practically contributing to achieving Coast Guard priorities. At the very least, embark on those activities that serve to ensure the Coast Guard does not suffer a deteriorating strategic capability as its stakeholders and adversaries become more expert in the digital domain.

3.1 Preface on Prior Findings: Deterrence and Current Coast Guard Practice

Deterrence is crucial to most U.S. Coast Guard mission areas. Coast Guard operations produce a deterrence effect. Yet deterrence has proven difficult to distinguish from other strategies, tactics, and techniques and has resisted meaningful measurement.

A 2012 technical report – *Deterrence and the United States Coast Guard: Enhancing Current Practice with Performance Measures* (UDI#1181) – found the U.S. Coast Guard engages in effective deterrence and related protection and prevention practices across several mission areas. Moreover, Coast Guard practices closely correspond with leading models of economic choice for which methods of performance measurement have been developed, vetted, and confirmed. The opportunity exists to enhance Coast Guard deterrence practice, especially, the measurement of deterrence by shaping Coast Guard deterrence strategies, data capture mechanisms, tactics, and techniques in accordance with these models of economic choice and other behavioral tools.

Deterrence is most likely to be achieved when several aspects of multi-dimensional presence are practiced within explicitly mapped social networks. Choosing when, where, and how to make investments in multi-dimensional presence can be substantively informed by findings related to economic choice. Tracking these investments and outcomes over time provides a means for measuring deterrence and driving improvement of deterrence. Real time data capture and improved command and control mechanisms will allow better measurement of the impact of resource allocation and deployment of assets in Coast Guard missions.

Deterrence is not currently differentiated from related strategies, tactics and techniques, such as surveillance, patrol, interdiction, disruption, inspection, community outreach, training, and more. Deterrence is not measured. Because deterrence is not differentiated or measured it is: not well-understood; is practiced sporadically; is disconnected from daily operations; and is resistant to consistent and rigorous improvement. Command and control tools and data capture practices are not optimized to scientifically measure the behavioral impact of Coast Guard choices. Especially in a resource-challenged environment featuring creative and adaptive potential adversaries and non-compliant stakeholders, there are benefits to practicing what can be measured (rather than investing in other more expensive strategies, tactics, and techniques).

Until quite recently the technical means to map social networks and apply multi-dimensional presence in a data-informed manner were not available. Practical technical means now exist and are worth examining in terms of readiness for measuring deterrence and driving improvement of deterrence.

3.2 Core Concepts of Data-Driven Deterrence

The emergence of Big Data and related analytics offers the United States Coast Guard a potentially powerful new tool with which to enhance safety, stewardship, and security in the maritime environment. As set out in a prior report[3], the Coast Guard has traditionally been a very effective practitioner of deterrence, operationally anticipating many of the most advanced academic theories of behavioral influence. This core competency emerged from an intimate relationship with maritime communities and has been based on long-time institutional experience with these populations' social dynamics, especially those of the fishing and shipping communities.

As the communities the Coast Guard serves increasingly migrate to a wide array of digitally-mediated social interactions and performance measures, (and as this digital domain expands to include populations not traditionally served by the Coast Guard), it is especially important that the Coast Guard develop sufficient digital intimacy to preserve its core competency in deterrence.

The data streams flowing out of (and into) maritime communities are analogous to the ports and channels (and backwaters) in which the Coast Guard and its predecessor organizations have long operated.

Deterrence displaces or delays unwanted behavior. More ambitiously, deterrence aims to permanently discourage an individual or individuals from such behavior. It operates through the prospect of pain or pleasure. Fear of future pain and hope of future pleasure influence choices, habits, and predilections. Individuals seek to avoid pain and maximize pleasure. Expectations of pain and pleasure vary, but given an individual's or group's specific sensibilities, a pattern of pain-minimization and pleasure-maximization tends to persist. An effective deterrence strategy rewards "good behavior" and punishes "bad behavior."

Such sanctions can be characterized as instrumental, social, and normative. Generally, instrumental efforts – such as patrols and inspections – generate structured data (e.g., time, place, and activity type) while social and normative efforts generate unstructured information (e.g., conversations, thoughts, and opinions).

3.3 USCG Implications: Seeding and collaborating with existing communities

Where the Coast Guard is most interwoven into a community it has the most options for engaging the community. When the Coast Guard understands how the community operates, what the community values, and the various levers of influence within a community and all its parts, the Coast Guard can mindfully choose how, when, and where to operate (or not operate) to support – and sometimes to shape – what the community values and how it sanctions what is valued. The more the Coast Guard is in a relationship with the community (ideally "of" the community), the more likely Coast Guard missions will be achieved.

[3] *Deterrence and the United States Coast Guard: Enhancing Current Practice with Performance Measures*, Research and Development Center, March 2012

The Coast Guard's ability to effectively and efficiently engage communities emerges from a humane tradition of encouraging safety and stewardship among a free people who are attached to the nation's waterways. Translating this tradition into the digital era is important. It will not be simple.

Respect is regularly referenced as a core-value of the Coast Guard. Especially in its relationship with maritime communities, this has generated significant strategic advantages. Respect extended is often returned. Collaboration and cooperation are cultivated. Retired Admiral James Loy and Donald Phillips offer this description of the Coast Guard approach:

> *Respect for the people you serve. Respect for your shipmates. Respect for every human being you may encounter inside or outside the organization, from bus drivers to admirals. There is also a deep-seated belief that all people have value, despite any differences they may have. It is essentially the embodiment of the Golden Rule: Treat people as you would have them treat you. But in the Coast Guard it is a hard and fast rule. Everyone must treat others fairly and with civility, consideration and dignity.*[4]

Dignity is closely associated (at least in many cultures) with privacy, autonomy, and the potential to exercise individual agency. In the pre-digital era, the Coast Guard has been remarkably effective in balancing its enforcement activities and its support activities. Compared with many enforcement agencies, the Coast Guard is perceived by communities it serves as much more than "cops in boats." The Coast Guard is cop and firefighter and lifeguard and customer and neighbor and often much more. The Coast Guard respects, even protects, the dignity of the communities it serves. In turn, it is respected. The respect extended to the Coast Guard is a significant force-multiplier earned over generations of service to maritime communities. Any loss of respect seriously degrades Coast Guard effectiveness.

The respect traditionally extended to the Coast Guard will not necessarily follow it into the digital domain and loss of respect in the digital domain could complicate Coast Guard activities in other domains.

Behaviors that generate particular respect in the digital domain include delivery of practical benefits, collaboration, and transparency (perhaps in that order). Google, for example, is perceived as providing its users with several practical benefits, many of which are generated by facilitating online collaboration. Public attitudes regarding Google's transparency are more ambiguous, but there seems to be sufficient collaborative benefit to outweigh what some call the "creepiness" factor involved in lack of transparency. As a government agency, the US Coast Guard will be expected to demonstrate much more transparency than Google, Facebook, Amazon or other strong players in the digital domain. Can the Coast Guard generate sufficient collaborative benefit?

The Coast Guard is involved in developing and delivering several resources that have potential social media applications. Just three examples:

- Nationwide Automatic Identification System (NAIS): Consists of approximately 200 VHF receiver sites located throughout the coastal continental United States, inland rivers, Alaska, Hawaii and Guam. NAIS collects AIS transmissions from local vessels. Currently, NAIS collects valuable maritime data in 58 critical ports throughout the United States for use by Coast Guard operators and port partners.

[4] Phillips, Donald T. and Loy, James M., *Character in Action: The US Coast Guard on Leadership*, Page 13

- Vessel Traffic Service (VTS): provides active monitoring and navigational advice for vessels in particularly confined and busy waterways. There are two main types of VTS, surveilled and non-surveilled. Surveilled systems consist of one or more land-based sensors (i.e., radar, AIS and closed circuit television sites), which output their signals to a central location where operators monitor and manage vessel traffic movement. Non-surveilled systems consist of one or more reporting points at which ships are required to report their identity, course, speed, and other data to the monitoring authority.
- Digital Selective Calling: Allows mariners to instantly send an automatically formatted distress alert to the Coast Guard or other rescue authority anywhere in the world. Digital selective calling also allows mariners to initiate or receive distress, urgency, safety and routine radio-telephone calls to (or from) any similarly-equipped vessel or shore station, without requiring either party to be near a radio loudspeaker.

None of these tools and services were originally conceived for the "digital domain." Each are based on pre-digital monitoring technologies more similar to the Internet-of-Things than social media. But these and other Coast Guard resources generate the potential of practical benefit around which social media strategies can be developed.

Others are, in fact, already accessing AIS-data to support social media strategies. Just one example: Global Fishing Watch, co-developed by Oceana, SkyTruth, and Google, deploys AIS archival data to generate a visualization of global fishing patterns. The visualization is then used to attract interest, generate discussion, influence norms, and prompt action by online and off-line communities.

As the team interviewed Coast Guard personnel and stakeholders, several examples, suggestions, and existing tools were referenced related to current and potential opportunities for the U.S. Coast Guard to engage the digital domain, but (at least at the operational level) no clear strategy could be discerned regarding the cumulative effects of digital engagements. They are more often conceived and deployed as narrowly tactical responses to specific problems.

The team perceives that the U.S. Coast Guard has existing practical benefits to offer the digital domain and a clear capacity to develop many more. What is missing, paradoxically, is a sustained process of collaborating with communities already active in the digital domain. We say this is paradoxical because so much of Coast Guard culture off-line is so intensely collaborative. Just as the Coast Guard proactively engages non-digital communities, there is a need for the Coast Guard to reach out to the growing digital community.

It could be very interesting (whatever the result) for the Coast Guard to transparently communicate an interest in collaborating with stakeholders across the maritime and digital domains to develop digital resources that would provide practical benefits and would also support the ability of the Coast Guard to better understand and effectively practice multi-dimensional (including digital) deterrence. It is possible that, no matter how benign the intention or communication, public responses would demonstrate that such actions by the Coast Guard would undermine public respect for the Coast Guard.

But if a significant cross-section of digital and maritime communities would welcome the opportunity to collaborate with the Coast Guard, this collaboration could: amplify use of already existing digital resources; identify opportunities for Coast Guard investment in new digital resources; position the Coast Guard as a

trusted neighbor in the digital domain; and generate strategic synergy in both the digital and non-digital domains.

3.4 Populations, Sub-populations and Social-identity Groups

> *"The real challenge isn't finding the needle in the haystack.*
> *It's finding and mastering the haystack itself."*
>
> H.O. Maycotte, *Forbes*, January, 2015

Epidemiology arose in the 19th Century as a statistical study of populations and sub-populations to identify sources and patterns of disease. Through a variety of population studies it has been possible to identify environmental and behavioral correlations that can guide actions designed to prevent and mitigate the outbreak of disease among populations. Similar statistically-informed strategies have increasingly been used in modern law enforcement to prevent and mitigate crime. Analogous techniques are deployed to manage a variety of contemporary networks, including the electrical grid, telecommunications systems, and supply chains.

The increasing availability of Big Data streams – both social and industrial – offer an unprecedented opportunity to situate an examined population's baseline conditions, fluctuations, possible sources of fluctuation, and opportunities to intervene to manage both baseline conditions and fluctuations. Correlation is, of course, different than causation. But correlations present decision-makers with helpful clues for closer examination and exploration.

Here is how the opportunity (and challenge) of Big Data was recently described by the Harvard School of Public Health:

> *Our ability to generate data has moved light-years ahead of where it was only a few years ago, and the amount of digital information now available to us is essentially unimaginable.*
>
> *"In the last five years, more scientific data has been generated than in the entire history of mankind," says Winston Hide, associate professor of bioinformatics at HSPH. "You can imagine what's going to happen in the next five." And this data isn't simply linear; genetics and proteomics, to name just two fields of study, generate high-dimensional data, which is fundamentally different in scale…*
>
> *In big data lies the potential for revolutionizing, well, everything. Police employing seismology-like data models can predict where crimes will occur and prevent them from happening. Astronomers using the Kepler telescope snag information on 200,000 stars every 30 seconds, which has led to the discovery of the first Earth-like planets outside our solar system. Businesses sifting social networking and supply-chain data dynamically tailor their products to fulfill desires we don't even know we have.*
>
> *The same phenomena are at play in public health. For some time, DNA sequencing has held big data's starring role—after all, a single human genome consists of some 3 billion base pairs of DNA.*

Researchers at HSPH and across the campus at Harvard are sequencing and analyzing human genomes to ferret out clues to infections, cancer, and noncommunicable diseases.

But the potential public health uses of big data extend well beyond genomics. Environmental scientists are capturing huge quantities of air quality data from polluted areas and attempting to match it with equally bulky health care datasets for insights into respiratory disease. Epidemiologists are gathering information on social and sexual networks to better pinpoint the spread of disease and even create early warning systems. Comparative-effectiveness researchers are combing government and clinical databases for proof of the best, most cost-effective treatments for hundreds of conditions—information that could transform health care policy. And disease researchers now have access to human genetic data and genomic databases of millions of bacteria—data they can combine to study treatment outcomes.

What is also emerging from the intersection of Big Data and public health is the potential to identify "nodes of influence" within a population that have an amplified effect on the whole set of population interactions. Here's one example reported by the Wharton healthcare management program:

Recent studies have found that one's chance of becoming obese increases by 57% if the person has a friend who became obese, and that one's chance of becoming obese is not only linked to the weight gain of one's friends, but also to the weight gain of friends of friends. Other research suggests that direct and even indirect social reinforcement can increase positive health behaviors. The implication is that if health care providers and officials can target "nodes of influence" within communities, they can more effectively improve outcomes and reduce the prevalence of the flu, sexually transmitted diseases or other conditions…"Social influence matters, and it matters a lot," says Dr. Larry Miller, CEO of MedNetworks. "If you can harness it, you are harnessing one of the most powerful forces that affect all of us, so why not do it." The company helps pharmaceutical companies and health care delivery companies understand how they can use social network technology to target nodes of influence in their physician or patient populations when designing promotions or wellness programs, with the goal of using resources more effectively and improving outcomes in patient populations.

There are several similar efforts underway outside public health. For example, the Federal Reserve Board is conducting sentiment analyses of social media networks to better understand the status and possible direction of financial markets. According to FedTech, the Fed's Chief Data Officer Micheline Casey explained,

The board's Division of Consumer and Community Affairs is starting to use sentiment analysis and social media to look at emerging areas and issues, related to the mortgage industry as well as problems with credit cards and auto loans… While social media data may not be as statistically accurate as information used for setting interest rates, she said, the sets will help to identify "indicators and signals that can give us better insight as to what's actually going on instead of just leveraging older data."

3.5 USCG Implications: Return on Influence

The 2012 phase of this study outlined a population-based approach to advancing these epidemiological principles by linking the maritime domain with the digital domain. The purpose was to generate data that will accurately inform decisions about where, when, and how to assert influence.

That previous study argued, "Deterrence is most likely to be achieved when several aspects of multi-dimensional presence are practiced within explicitly mapped social networks. Choosing when, where, and how to make investments in multi-dimensional presence can be substantively informed by findings related to economic choice. Tracking these investments and outcomes over time provides a means for measuring deterrence and driving improvement of deterrence. Real time data capture and improved command and control mechanisms allow better measurement of the impact of resource allocation and deployment of assets in Coast Guard missions."

The Coast Guard can achieve this insight by mapping the role and character of Social Identity Groups (SIGs) relevant to Coast Guard missions and accurately modeling how SIGs are both the source of and influenced by Instrumental, Social and Normative (ISN) factors. In this way the Coast Guard can deploy ISN interventions that are statistically most likely to enhance deterrence.

The ISN framework is one way to express the range of influence vectors operating in any decision space. **Instrumental factors** include law enforcement actions, economic outcomes, and physical impediments. **Social factors** include attitudes, sentiments, and actions by other humans. **Normative factors** include perceptions of fairness, legitimacy, reciprocity and right or wrong. In many cases ISN factors are as powerful in their indirect influence as their direct influence.

Non-compliance with laws and regulations can be understood (in an epidemiological context) as contagious disease. The goal is to contain the contagion. Non-compliance undermines the safety, sustainability, and fairness of crucial aspects of maritime trade and resources. Actively enforcing compliance when deterrence fails can be operationally complex and expensive. Effective and especially *cost-effective* enforcement depends on deterrence. Only when the vast majority of maritime actors are inclined to cooperate with the Coast Guard can enforcement resources hope to be effectively deployed. Only when non-cooperation is limited to a distinct minority of maritime actors can the limited resources of the Coast Guard effectively target non-compliant and adversarial behavior.

Better understanding the social and normative characteristics of non-cooperation will enable the instrumental resources of the Coast Guard to be more efficiently utilized. This understanding should also expose how the social and normative characteristics of the whole population can be encouraged to influence and reverse early tendencies toward non-cooperation. Adversaries seek to disrupt or destroy for purposes contrary to the national interest and the shared interest of most maritime actors. When Coast Guard deterrence strategies, tactics, and techniques prevent non-compliance or adversary action the national interest is advanced.

A Deterrence Impact Modeling Environment has been proposed as a strategic system and management tool with which Coast Guard decision-makers can enhance prevention, protection, and compliance by selecting strategies, tactics, techniques and procedures that increase psychological and social influences to discourage unwanted behavior and encourage desired behavior and in this way increase the likelihood of deterrent effects.

3.6 Common-pool-resources, Economic Choice, and Maritime Communities

"Given enough digital records and enough computing power, a new vantage point on human culture becomes possible…"

Erez Aiden and Jean-Baptiste Michel
Uncharted: Big Data as a Lens on Human Culture

The United States Coast Guard advances stewardship, safety, and security in the maritime environment. Each of these strategic objectives are in the language of economists, "common pool resources." This is a largely uncontroversial claim in terms of the Coast Guard stewardship mission. If it is also true for safety and security, several novel strategic implications follow.

A common-pool-resource is a "class of goods or events" for which it is difficult to exclude individuals from accessing and when claimed by one individual serves to subtract from potential benefits available to other individuals. The classic example is when one fisherman lands a ton of fish those fish are not available to other fishermen. [5]

Common-pool-resources (CPRs) are characterized by lower levels of exclusion and higher levels of subtractability.

[5] Ostrom, Elinor, Gardner, Roy and Walker, James "*Rules, Games, and Common-Pool-Resources"*, University of Michigan Press, 1994

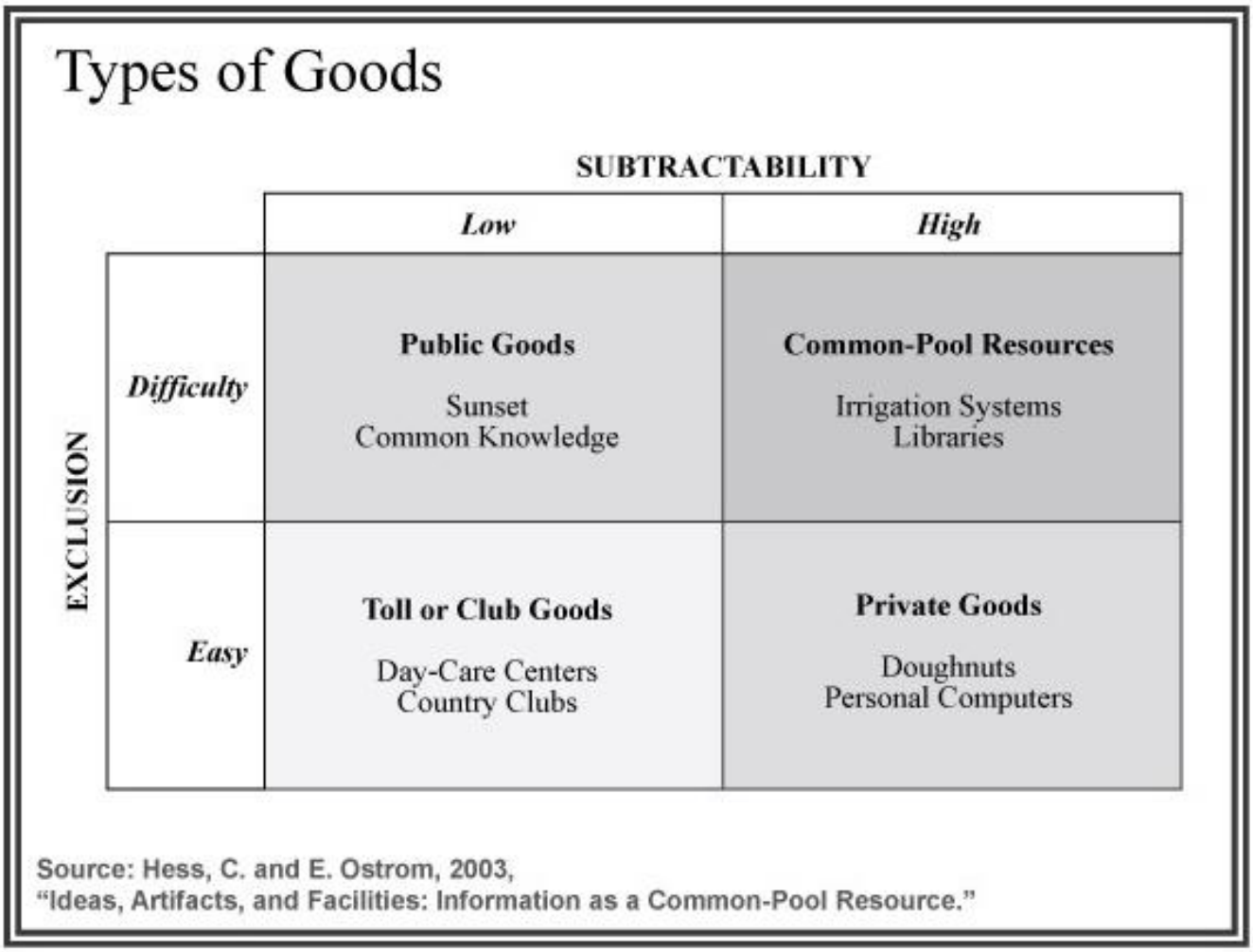


Figure 2. Exclusion/subtractability chart for different classes of resources.

Safety and security have traditionally been conceived as public goods: difficult to exclude, but having little or no subtractability. Increased safety and/or security for one individual, group, place, or time has not been understood to reduce the benefit for other individuals, groups, places, or times.

If this was mostly true in the past, it is less-and-less true in many contemporary contexts. In terms of security, it is increasingly recognized that success with one terrorist group, one cartel, one human-trafficking organization, one maritime zone mostly displaces rather than reduces the threat. Even in terms of safety, attention given to one concern (particularly attention given by an "outside" authority) can actually increase other safety problems. As risk increases (especially as the population involved in the maritime environment increases), the team perceives that safety and security will often join stewardship in exhibiting increased subtractability.

To the extent increased subtractability can be confirmed in many safety and security contexts, then more attention should be given to evidence that what influences human behavior regarding common-pool-resources is considerably different than what influences behavior for public, private, or toll goods.

As previously noted in this report, it is possible to characterize Coast Guard activities as having Instrumental, Social or Normative influence. Instrumental means are especially effective when exclusion from the class of good or event is comparatively easy. Protective or preventive actions are specifically designed to exclude. But when exclusion is difficult, Social and Normative sources of influence become considerably more important. What means are most effective to individuals and groups choosing to exclude themselves (and actively monitor and intervene to ensure other individuals and groups make a similar choice)?

Over the last thirty-plus years Dr. Elinor Ostrom and other scholars have conducted extensive field experiments to answer this question. Here is how Dr. Ostrom explained their results in her 2009 Nobel Prize lecture:

> *While it is not possible yet to point to a single theory of human behavior that has been successfully formulated and tested in a variety of settings, scholars are currently positing and testing assumptions that are likely to be at the core of future developments. These related to (1) the capability of boundedly rational individuals to learn fuller and more reliable information in repeated situations when reliable feedback is present, (2) the use of heuristics in making daily decisions, and (3) the preferences related to benefits for others…*
>
> *The… variables that are most important differ depending on which interactions (such as monitoring, conflict, lobbying, and self-organization) or longer-term outcomes (such as over-harvesting, regeneration of biodiversity, resilience of an ecological system to human nature-induced disturbances) one wishes to predict. A set of ten variables have been identified across many field studies as impacting the likelihood of users self-organizing in order to overcome a common-pool-resource dilemma…*

In our examination of Coast Guard practice, the team has observed policy, strategy, tactics, techniques, and procedures (especially in regard to the stewardship mission) that are entirely consistent with Ostrom's findings. But in terms of safety and security missions, decisions made and actions taken often seem to presume a potential for exclusion that is not always realistic.

In its stewardship mission the United States Coast Guard is expert at deploying a range of interventions that serve to increase the likelihood of individuals and communities choosing to take (or refrain from) actions that enhance the long-term health of a common-pool-resource. Can similar expertise be deployed for safety and security?

Ostrom and others have derived from their field research ten variables that appear to have an amplified impact on whether or not a group self-organizes to enhance sustainability of the common-pool-resource. The following attempts to "translate" these variables using terms more common to security and safety:

1) *Size of System:* Large enough to be economically, politically or socially persistent, small enough to be meaningfully monitored. Probably tied to specific port communities.
2) *System Productivity*: Community recognition of emerging scarcity or potential increase in subtractability. Is there a recognized safety or security problem?

3) *Predictability of System Dynamics*: More rather than less predictability in terms of how behavior adjustment will impact system. Is the community convinced something practical can be done to enhance safety and security?
4) *Resource Mobility*: Stationery resources – or specific places – are more favorable to self-organization than resources that are easily mobile. This variable is more supportive of safety and security than is often the case with stewardship.
5) *Number of Users*: Smaller is typically better, larger groups can complicate processes. But depending on the size and dynamics of a system more users may facilitate monitoring and enforcement.
6) *Leadership/entrepreneurship*: Preexisting organization and effectiveness tends to support progress in engaging novel challenges. Success in stewardship should set the stage for safety and security.
7) *Norms/Social Capital*: Preexisting and widely shared moral and ethical standards tend to support self-organization and effectiveness.
8) *Knowledge of SES/Mental Models*: "When users share common knowledge of… how their actions affect each other… they will perceive lower costs of organizing."
9) *Importance of Resource*: If safety and/or security is perceived by the community as important, it is more likely the community will be willing to self-organize to sustain the resources.
10) *Collective Choice Rules*: "When users… have full autonomy at the collective-choice level to craft and enforce some of their own rules, they face lower transaction costs as well as lower costs in defending a resource against invasion by others."

Big Data with emerging analytical methods presents the Coast Guard with an ability to understand and track these variables in a way never before possible. This understanding and tracking can inform decisions that serve to deter unwanted behavior and encourage stewardship, safety and security of the maritime domain.

3.7 USCG Implications: Empirically-based Methods

The Coast Guard needs to clarify the structure of communities and the social, technical, economic, and ecological systems on which communities depend. This enhanced understanding will allow much more accurate targeting of particular Coast Guard interventions. Further, interventions may not work the same way over time. As structural variables change, the Coast Guard needs to have ways of learning and adapting to these changes.[6]

Complex systems are partially decomposable in their structure. Three aspects of decomposability of complex subsystems are important to achieve a better understanding of complex communities and their networks in order to craft ways to improve them. The first aspect is the conceptual partitioning of variables into classes and subclasses. The second aspect is the existence of relatively separable subsystems that are independent of each other in the accomplishment of many functions and development but eventually affect each other's performance. The third aspect is that complex systems are greater than the sum of their parts.

At the broadest conceptual level, one can develop a general framework – a conceptual map – that can be used as the starting point for conducting the study of systems that underlie communities. Figure 3 below presents a simple, very general framework for the highest-tier variables that must be analyzed when examining linked systems. At this broad level, one can begin to organize an analysis of how attributes of (i)

[6] Ostrom, Elinor, "A Diagnostic Approach for Going beyond Panaceas", *Proceedings of the National Academy of Sciences*, Volume 4, Number 39, September 25, 2007.

a resource system (e.g., fishery, port, supply chain, area), (ii) the resource units generated by that system (e.g., fish, water, oil, containers), (iii) the users of that system, and (iv) the governance system jointly affect and are indirectly affected by interactions and resulting outcomes achieved at a particular time and place.

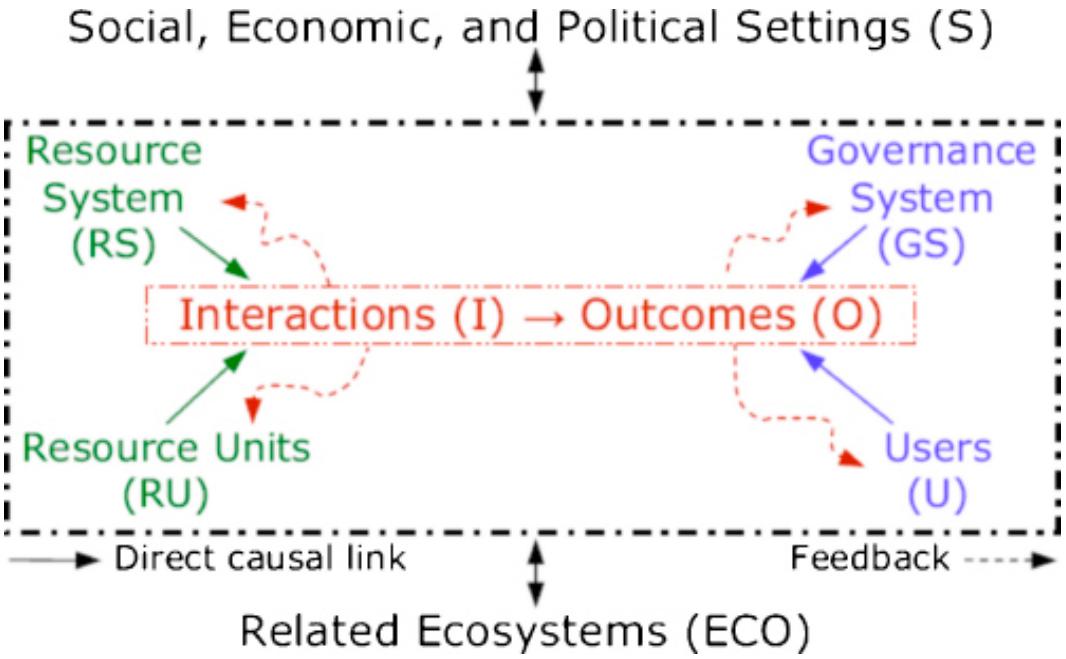


Figure 3. A multitier framework for analyzing community systems.

Using such a framework enables one to organize how these attributes may affect and be affected by the larger socioeconomic, political, technical and ecological settings in which they are embedded, as well as smaller ones.

Because this is a decomposable system, each of the highest-tier conceptual variables in Figure 4 can be unpacked and related to other unpacked variables in testable theories relating the outcomes of human behavior with a community and its systems. Figure 4 below lists major second-tier variables that have been shown in empirical studies to impact diverse interactions and outcomes. They are the initial core of conceptual variables needed to identify the broad types of systems operating at a particular location in time and space so an accurate diagnosis of the reasons for desired or undesirable outcomes can be identified. In addition to the broad second-tier variables identified in Figure 4, many more specific variables are identifiable at deeper levels. Research has been underway for several years to develop this diagnostic framework further and link it to rigorous empirical research findings. A major challenge is defining all variables so the conceptual logic of linking more specific concepts to more general concepts is clear and open to further discourse and development.

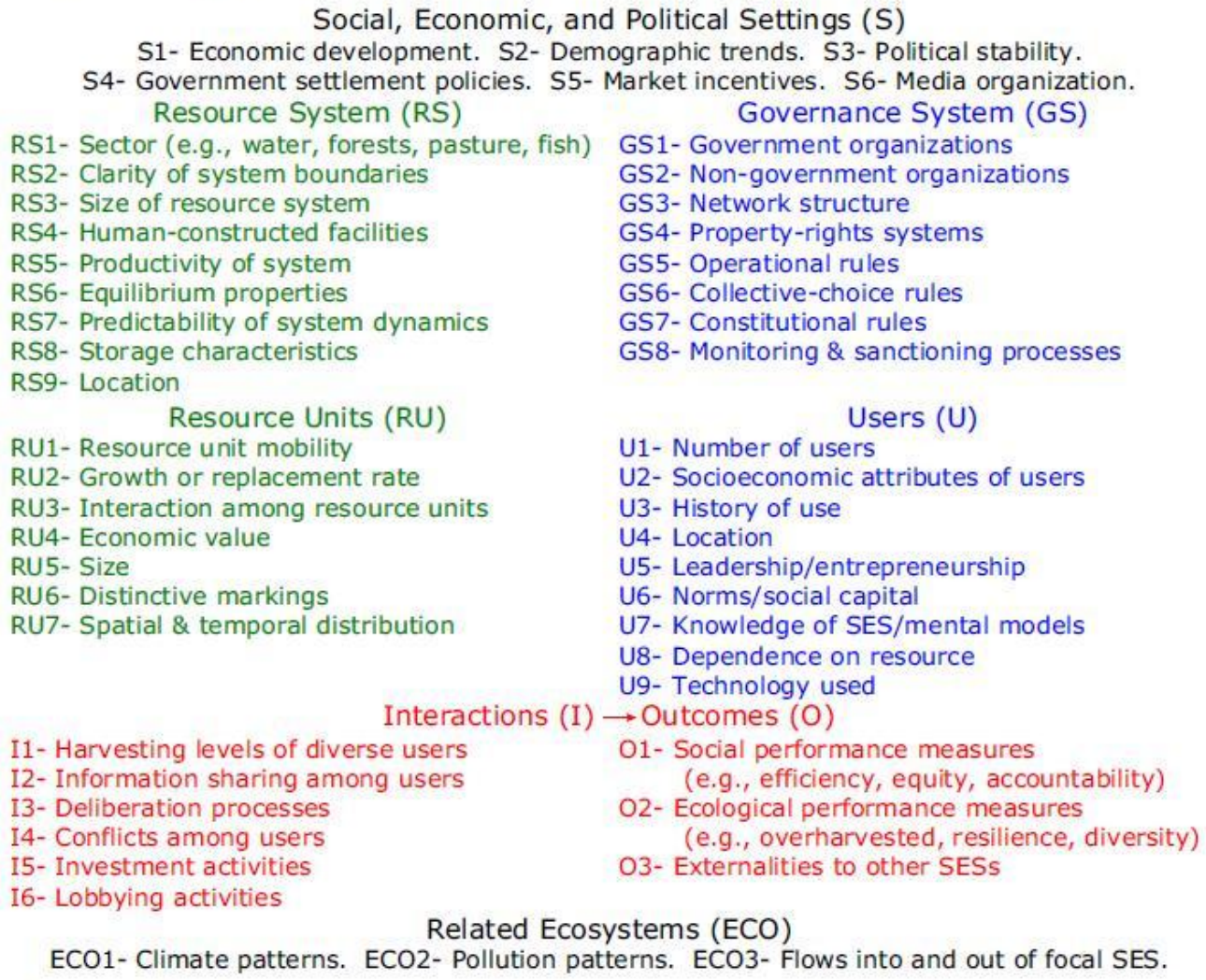


Figure 4. Second-tier variables in framework for analyzing an SES.

The long-term goal for the US Coast Guard is to recognize which combination of variables tends to lead to relatively sustainable and productive use of the maritime environment. This presents compound puzzles nested in compound puzzles. The key is assessing which variables at multiple tiers across the biophysical and social domains affect human behavior and outcomes over time. Big Data with analytics (drawing on archival and real-time data streams) provides a practical opportunity to do this that will build on itself. Each success and failure will enhance the probability of the next targeted intervention.

3.8 Data Exhausts and Navigating the Digital-behavioral Interface

"Data exhaust"—the trail of clicks that internet users leave behind from which value can be extracted—is becoming a mainstay of the internet economy.

The Economist, Data Data Everywhere

It is worth stating again that as recently as five years ago, the team would not have recommended that the US Coast Guard attempt to undertake what is outlined in the previous section. Even today the processes are pioneering. But pioneers may generate long-term (even self-optimizing) benefits precisely because they are pioneers.

Research undertaken over the last year suggests to the team that the maritime communities which the US Coast Guard serves are mostly *not* early adopters of social media or the Internet-of-Things (IoT).[7] We are not the only ones to reach this conclusion. In a special report on the IoT and shipping, an editorial in the May 2014 edition of *Futurenautics* offered,

> *Manufacturers, shipping's customers, are beginning to wrestle with a whole new range of technological and strategic questions. Questions which shipping should at least understand, and begin to actively help them solve. But a brief survey of some of our big customers sends a stark message: shipping is already the weak link in many operations… Shipping has been described as operating in the stone age, and when one compares the level of connectivity, digital operations, insight, data and intelligence involved in our customers' businesses, as compared to the average ship operator it's hard to dispute that. But we do have the opportunity to change. Maritime connectivity has advanced massively, and that opens the doors to closer digital integration and adding value of which shipping hasn't always been capable in the past.*

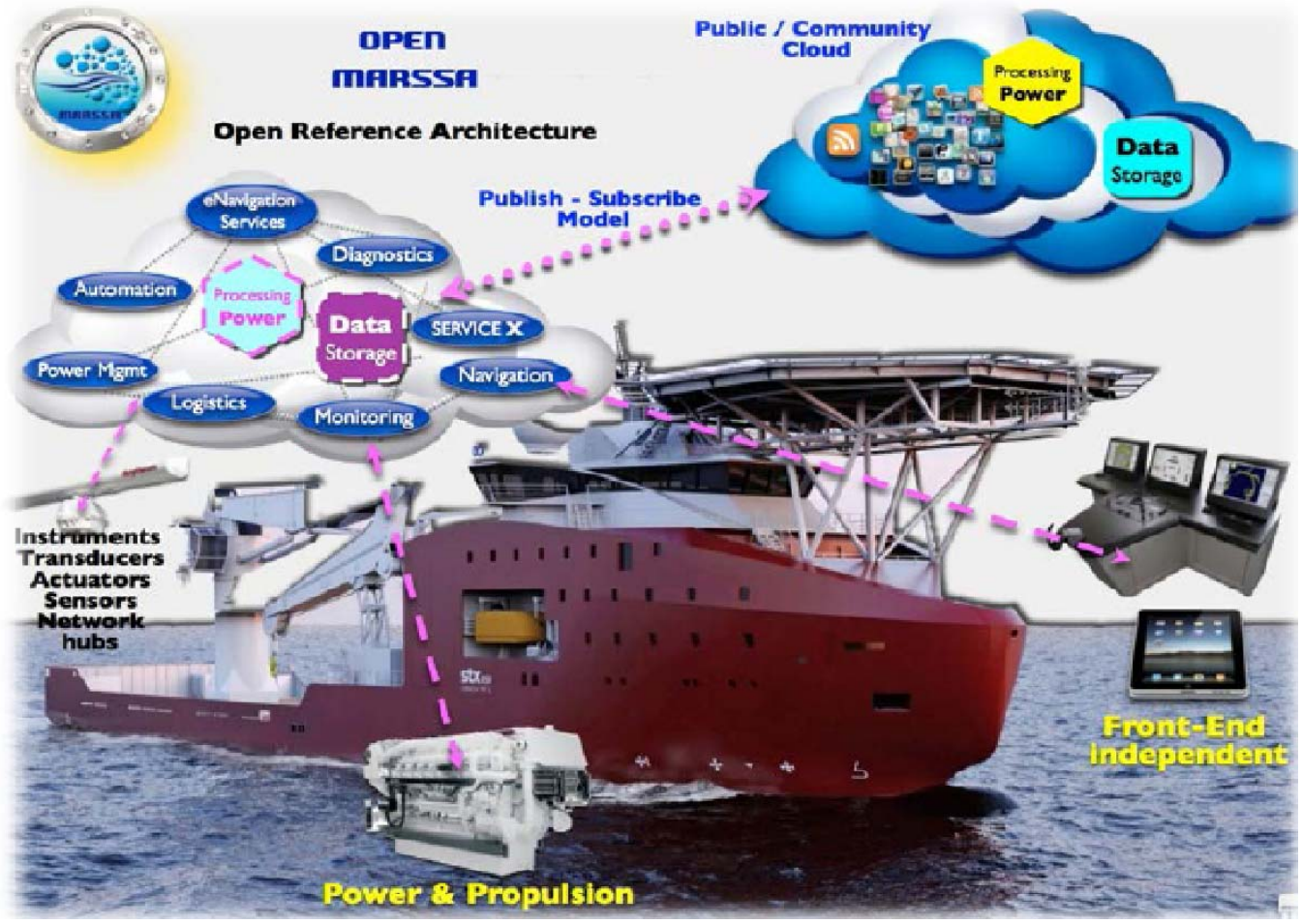
We would argue the current and coming shift is the result not just of connectivity but also available computing-power. In any case, the same developments that will now allow the US Coast Guard to engage Big Data analytics are *compelling* many of the Coast Guard's stakeholders to move much more aggressively in application of large scale employment of digital data.

The increasing number of participants and institutional players (and the pace of adoption) can already be discerned. For example,

- At the January 2015 Consumer Electronics Show, Swedish telecommunications vendor Ericsson launched its Maritime ICT Cloud designed to deliver a full-range of digital connectivity to the maritime community.
- Kongsberg Maritime is already applying the IoT to deliver systems for dynamic positioning and navigation, marine automation, safety management, cargo handling, subsea survey, and satellite positioning.
- The San Francisco start-up, Spire, claims that its satellite based monitoring of maritime-oriented IoT will help combat illegal fishing, capture pirates, support search and rescue, and monitor trade.

What most of these applications have in common is the ability to derive a wide array of situational awareness inputs from otherwise prosaic data streams: radio transmission (or their absence), engine telemetry, digital uploads of fuel usage, and such can also be used to construct an understanding of the maritime operating environment for an individual ship.

[7] Recreational boaters are active users of social media. Ports are increasingly using the IoTs. But our research suggests – the research was not sufficiently deep to be sure – that fishers and shippers have been slow to adopt.

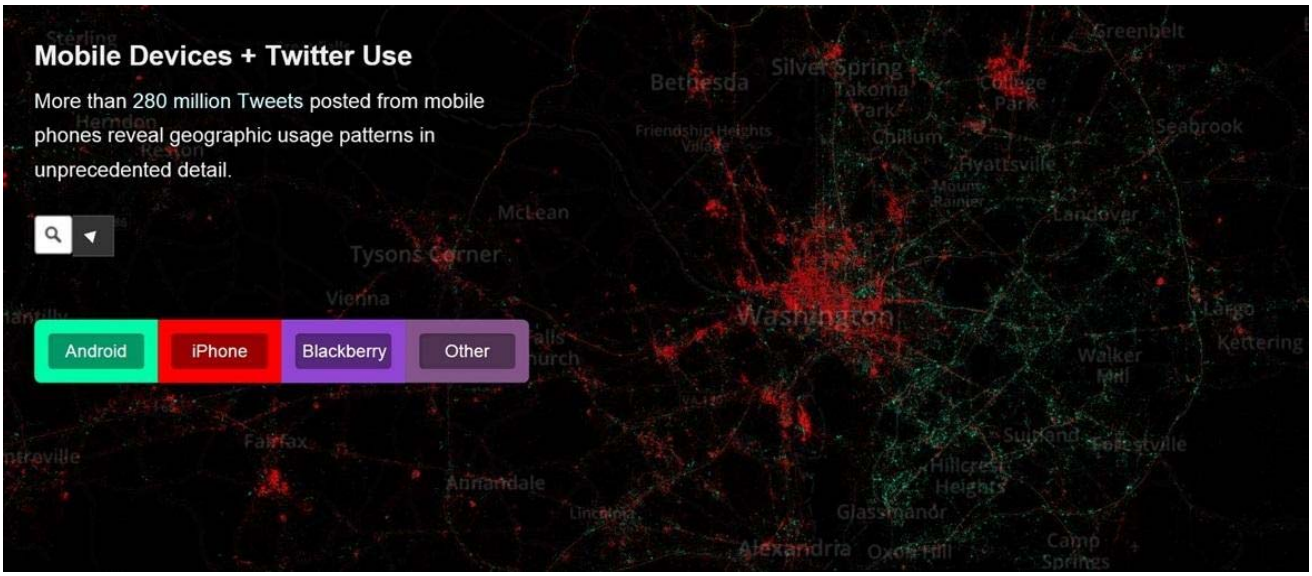


Source: *Geir Fagerhus, MARSEC-XL International*

Figure 5. New forms of data generation and collection in the shipping industry.

A collection of digital signals (Figure 5) for an individual ship can be used in such a way to observe the ship in a manner never anticipated in the original purpose of any single signal. This wider angle on reality is even more powerful when the number of those sending signals is multiplied.

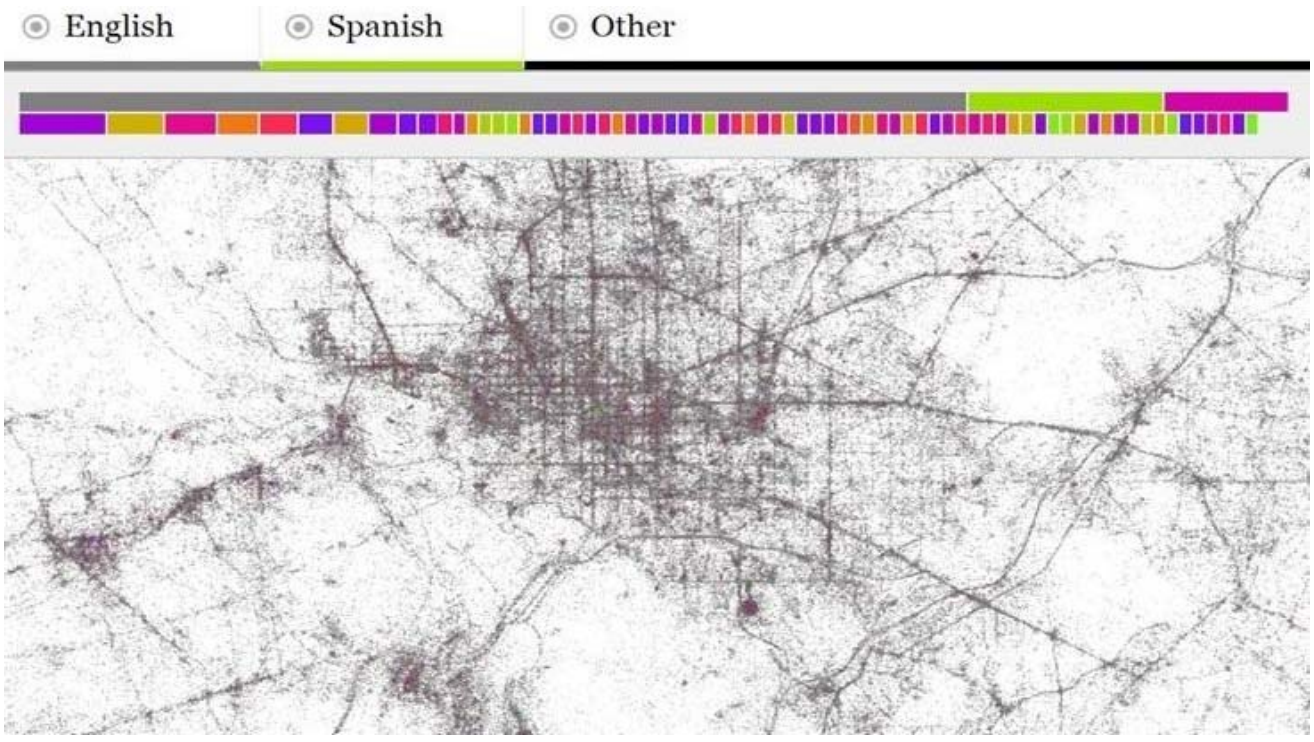
For example in Figure 6 below, twitter-feeds have been sorted by sending device. Each brand of phone is a different color. What this has exposed again and again, in city after city, is that device locations often reflect economic stratification. For example iPhones, in red, are predominantly in wealthy sections of the city while Android phones, in green, have more coverage in poorer sections.



Original at MapBox

Figure 6. Twitter users mapped according to the input device they are using.

In Figure 7, twitter-feeds are sorted by language used: English is gray, Spanish is green and all other languages (in this version) are magenta. It is worth noting that in neither of the visual displays shown on this page has a cartographic map been used. Rather, the display of collected geo-location data "creates" a map.



Original at MapBox

Figure 7. Twitter feeds sorted by language.

Using much smaller sample sizes than those displayed on the prior page, the team explored whether maritime communities (commercial fishing communities, in this case) could be distinguished within the whole population of twitter users. Theoretically, specific SIGs are distinguishable by, among other variables, the lexicon they use. The appearance of Statistically Improbable Phrases (SIPS) is a potential means of identifying and better understanding SIGS. For additional information on SIGs, see Appendix B.

Given time available and other resource limitations, this process of exploration began with a collection of five twitter accounts that had close relationships with commercial fishing. Word frequencies were then generated for an aggregation that extended over several weeks. The word frequencies for this aggregation were then compared to average word frequencies in the 5000 word Corpus of Contemporary American English (COCA) created by Dr. Mark Davies at Brigham Young University.[8] As the plot below indicates, these aggregations did display clear variations from the mean. The research process and sample size involved does not support definitive conclusions. But these early findings are suggestive that it is already possible to differentiate specific maritime communities from the vast ocean of social media data.

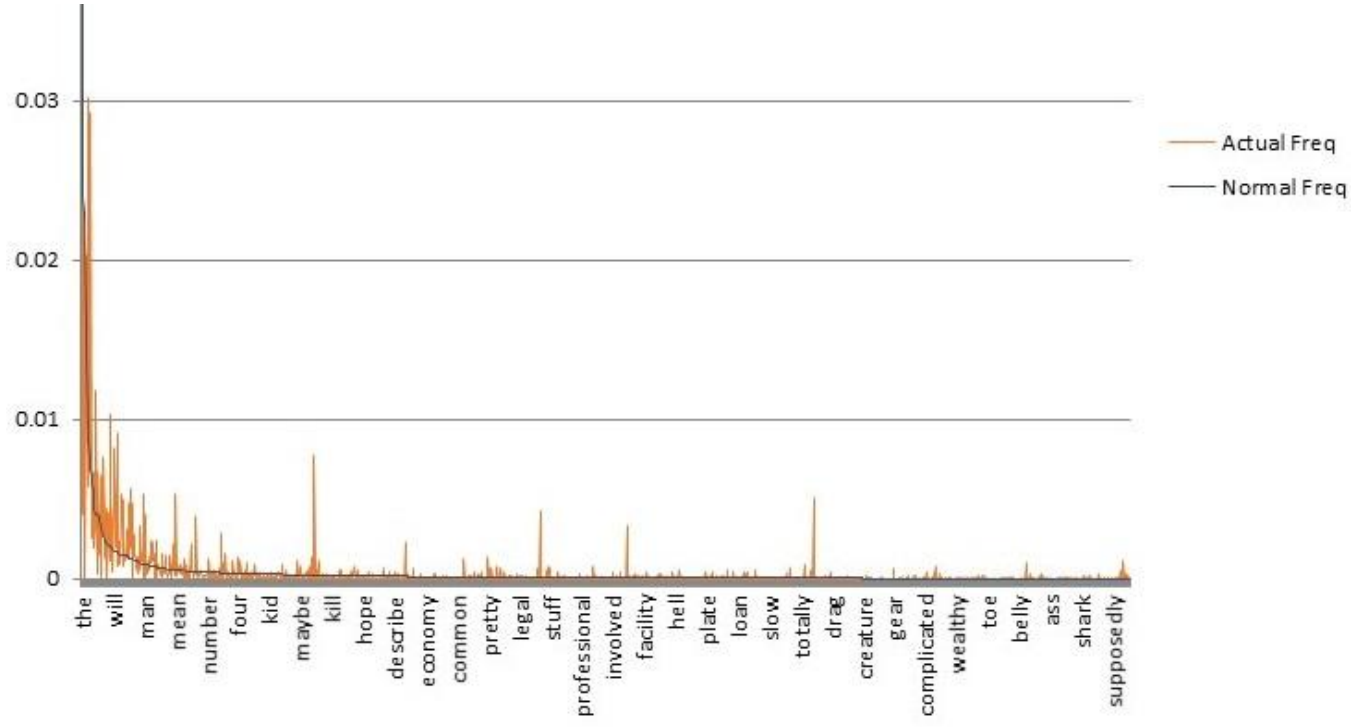


Figure 8. SIG speech patterns emerging from normalized English lexicon.

3.9 USCG Implications: Visualizing Vectors of Influence

To date, the variables active in community structures have been painstakingly derived from longitudinal studies and game-system models. This prior work has been rigorous and deeply informative, but innately limited.

[8] The Corpus of Contemporary American English now includes over 450 million words and can support much more detailed analysis than was undertaken for this initial exploration.

Elinor Ostrom closed her 2009 Nobel Lecture as follows:

> *To explain the world of interactions and outcomes occurring at multiple levels, we also have to be willing to deal with complexity instead of rejecting it. Some mathematical models are very useful for explaining outcomes in particular settings. We should continue to use simple models where they capture enough of the core underlying structure and incentives that they usefully predict outcomes. When the world we are trying to explain and improve, however, is not well described by a simple model, we must continue to improve our frameworks and theories so as to be able to understand complexity and not simply reject it.*

Today – in a way simply not possible in 2009 – we have the ability to deal directly and in almost real-time with complexity; not just in the form of a model, but in the actual unfolding of human/machine behavior and interaction. The rapidly rising ocean of digital data offers an unprecedented opportunity to engage reality with a fidelity that will increase as the data (and our engagement with it) accumulates.

Research findings (and commercial development) over the last two years have only reinforced the 2012 recommendation that the Coast Guard move toward development of Data Driven Deterrence (D^3). Findings and commercial progress have also confirmed the operational potential of a Deterrence Impact Modeling Environment (DIME) that, consistent with the prior work of Ostrom and many others would allow Coast Guard decision-makers to make choices that reflect instrumental, social and normative factors emerging in real-time within particular maritime environments.

DIME would enable a spectrum of engagement between the Coast Guard and various SIGs (and between SIGs, some of which have no direct contact with the Coast Guard). This spectrum of engagement ranges from positive to negative and involves instrumental, social, and normative interventions. Full implementation of DIME would facilitate the Coast Guard's ability to apply and modulate a mix of interventions to:

- Optimize the direct influence Coast Guard action has on the behavior of individual SIGs.
- Optimize how the totality of SIGs in a particular community (e.g., port or fishing grounds or waterway) interact with individual SIGs and enhance the indirect influence of the Coast Guard through these networks.
- Mobilize more "positive" SIGs to assist the Coast Guard in deterring more "negative" SIGs.

This would be achieved by engaging the SIGs with multidimensional ISN interventions as visualized in Figure 9 and Figure 10. Based on "Big Data" analysis of available information applied using cooperative game theory and other models, DIME provides Coast Guard decision makers with an assessment regarding which mix (and often in what sequence) of ISN interventions is most likely to produce desirable results.

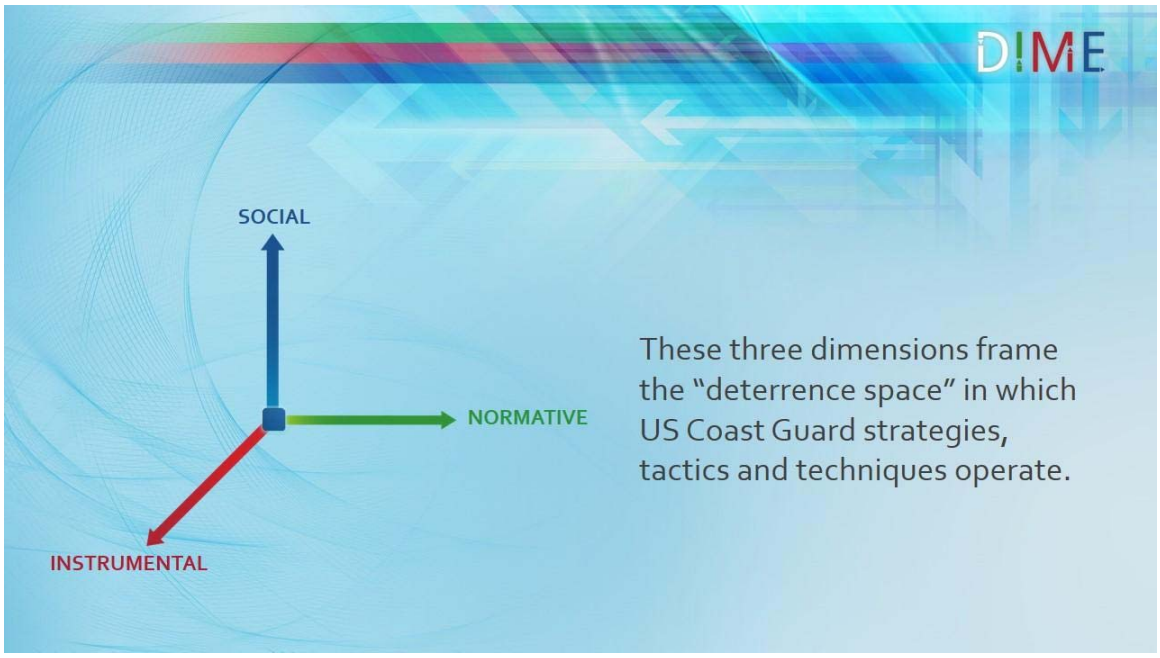


Figure 9. The instrumental, social and normative space.

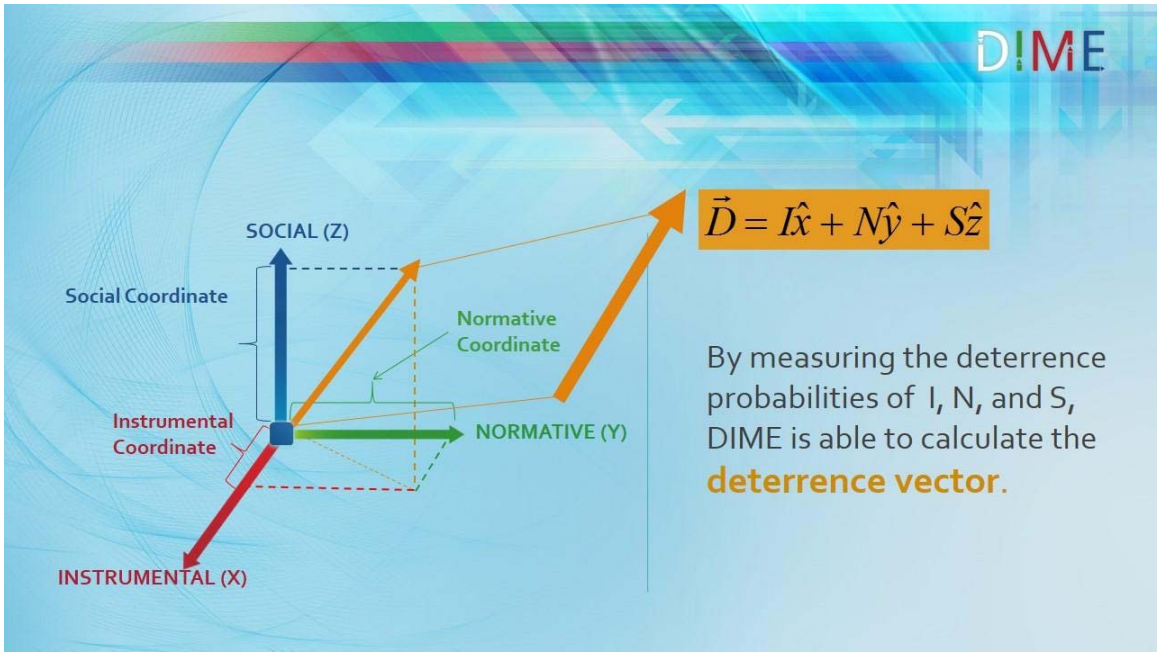


Figure 10. The deterrence probability vector.

The Coast Guard is (and always has been) a kind of apothecary, a pharmacist mixing various ingredients, mild or strong, to produce a healthy outcome. The effectiveness of the prescription has depended on experience, judgment, and luck. **Experience and judgment will still be needed. DIME will display probabilities, not certainties.** DIME can reduce the role of luck by increasing the role of real-time data processed specifically to inform the experience and judgment of Coast Guard personnel.

Over time, DIME and other forms of data-driven deterrence will expose whole new pharmacopoeias of potential interventions. To start, there are two rich sources of effective (or ineffective, even deadly) interventions: incentives and disincentives. DIME can be used to helpfully measure combinations of each.

3.10 Incentives, Disincentives and Reinforcing Social/Normative Influence

> *"Individual behavior is strongly affected by the context in which interactions take place rather than being simply a result of individual differences."*
>
> Elinor Ostrom, *Nobel Prize Lecture*

Deterrence is differentiated from other strategies by an intention to influence human motivation. In its most ambitious expressions, deterrence shapes fundamental categories of choice. Effective deterrence creates a context where certain desires and behaviors are reinforced and other behaviors are essentially framed as being beyond serious consideration, even if the desire persists.

The framework shown in Figure 11 reflects decades of field research. Attributes (including instrumental interventions) and rules (including social and normative variables) combine to shape interactions of actors with each other and their situations to produce a variety of outcomes. By shaping attributes of the physical world (for example, through persistent protection) and/or the attributes of community (for example, through active voluntary monitoring) and/or the rules-in-use (for example, non-fishing zones and seasons). In many ways, this is just an organized abstraction of common sense.

But the emergence of increasingly ubiquitous social media and the Internet-of-Things, combined with the potential of Big Data analytics, will now allow "common sense" to draw on data driven analysis of the relationships between these attributes, rules, and the action arena. This will allow those with access to and understanding of the data to anticipate how tweaks in one attribute or rule are likely to influence patterns of interaction and, therefore, outcomes.

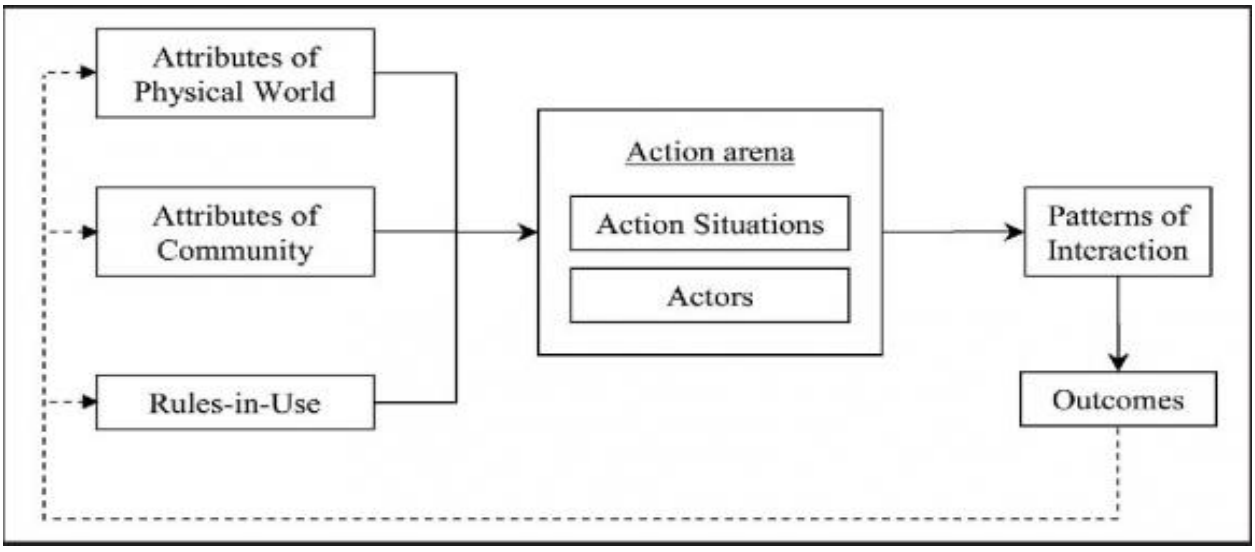


Figure 11. A framework for institutional analysis, Ostrom et al.

Since the 1990s, sustained attention has been given to how the framework outlined above can be expressed using Game Theory. The so-called *Prisoner's Dilemma* is probably the best known example of Game Theory. A relationship exists where a decision by one party can have a profound effect on a specific other. The full range of decision options is known. The optimal solutions for each party are known. The parties are unable to communicate with each other. Game theory has demonstrated a considerable ability to project the probable choices of the parties.

Different relationships and outcomes emerge when the full range of decision options is not known and/or the optimal solution for each party is not known. This is Bounded Game Theory, where the full scope of rationality is "bounded" or beyond knowing. But the mathematics with which Game Theory can be expressed still demonstrates an ability to calculate the range of choice, the matrix of influences behind such choices, and even an individual's optimal choice, given other choices. The calculation of a Nash Equilibrium, for example, can be a powerful tool for understanding any set of relationships.

Cooperative Game Theory (CGT) – a bounded environment featuring the ability of participants to communicate – was a particular interest of Elinor Ostrom and continues to be the object of significant experimentation. CGT is arguably the version most consistent with most realities, where interactions and consequences are very complicated or even complex, but participants are able to communicate and otherwise engage each other across a spectrum of relationships and potential interventions.

Multiple and diverse research efforts have found that CGT is especially adept at modeling and anticipating the probabilities of human choice when six variables are present to some significant degree:

1) Communication is feasible with the full set of participants. When face-to-face communication is possible, participants use facial expressions, physical actions, and the way that words are expressed to judge the trustworthiness of the others involved.
2) Reputations of participants are known. Knowing the past history of other participants, who may not be personally known prior to inter action, increases the likelihood of cooperation.
3) High marginal per capita return (MPCR). When MPCR is high, each participant can know that their own contributions make a bigger difference than with low MPCR and that others are more likely to recognize this relationship.
4) Entry or exit capabilities. If participants can exit a situation at low cost, this gives them an opportunity not to be a sucker and others can recognize that cooperators may leave (and enter other situations) if their cooperation is not reciprocated.
5) Longer time horizon. Participants can anticipate that more could be earned through cooperation over a long time period versus a short time.
6) Agreed-upon sanctioning capabilities. While external sanctions or imposed sanctioning systems may reduce cooperation, when participants themselves agree to a sanctioning system they frequently do not need to use sanctions at a high volume and net benefits can be improved substantially.[9]

In most "action arenas" involving stewardship and safety, these six variables are present. In many action arenas involving the security mission these variables are (or could be made to be) present.

[9] Ostrom, Elinor, Beyond Markets and State: Polycentric Governance of Complex Economic Systems (Nobel Prize Lecture) December 2009

Up to and until today, most research in CGT has depended on rigorous observational field research with retrospective analysis or simulated laboratory experiments. The increasing availability of data-streams associated with social media and the IoT generates the potential to integrate CGT with real-time or near real-time human communication and behavior.

If the Coast Guard would decide to undertake further exploration at the intersection of CGT and Big Data, it would be a pioneer, but it would not be alone. This analytical approach is already being applied to issues of national security, human evolution, and many other areas. We have included in an appendix a recent application to identify the sources of unintended discrimination in digital commerce. The math is known. The math requires thoughtful customization to different action arenas.

Using CGT to model the interactions (the influence vectors) among key populations will generate for Coast Guard decision-makers a data-driven set of probabilities that can, along with the experience and judgment of the decision-makers, help determine what combinations of action and non-action by the Coast Guard are most likely to generate behaviors which advance the Coast Guard Mission.

In her Nobel Prize Lecture, Elinor Ostrom reflected on a lifetime of professional engagement focused on how individuals and communities, official and non-official parties, good and bad actors – and many other dichotomous variables – interact to generate shared outcomes. She concluded:

> *The most important lesson for public policy analysis derived from the intellectual journey I have outlined here is that humans have a more complex motivational structure and more capability to solve social dilemmas than posited in earlier rational-choice theory. Designing institutions to force (or nudge) entirely self-interested individuals to achieve better outcomes has been the major goal posited by policy analysts for governments to accomplish for much of the past half century. Extensive empirical research leads me to argue that instead, a core goal of public policy should be to facilitate the development of institutions that bring out the best in humans. We need to ask how diverse polycentric institutions help or hinder the innovativeness, learning, adapting, trustworthiness, levels of cooperation of participants, and the achievement of more effective, equitable, and sustainable outcomes at multiple scales.*

3.11 USCG Implications: Full-spectrum Deterrence

Prior research has indicated that "multi-dimensional presence" is a key strategic differentiator (often a force multiplier) for the US Coast Guard. In distinct contrast to many other types of government intervention, the US Coast Guard already combines instrumental, social, and normative factors. This approach especially characterizes execution of the stewardship and safety missions. Experienced Coast Guard officers are quite adept at considering how the existing values and social relationships within a commercial fishing port, for example, can be leveraged to minimize the need for harsh enforcement actions. The Coast Guard has traditionally practiced what Elinor Ostrom and her colleagues have empirically demonstrated.

This current capacity is reflected in Coast Guard policy, doctrine, training, and operations. But it is an approach that has been practically impossible to measure and, as a result, sometimes difficult to defend. In fact, the need to measure performance has probably tended to privilege instrumental interventions over social and normative. Outputs are usually easier to measure than outcomes, even to the point that output measures can sometimes distract from achieving meaningful outcomes.

But today the online social interactions of Coast Guard stakeholders can be observed and measured. The relationship of these social interactions to actual behavior can be observed and measured using data-streams emerging from the Internet-of-Things. The normative values and perceptions of stakeholders can be observed and measured using sentiment analysis and other statistical tools. The relationship of normative values and perceptions to actual behavior can be observed and measured. The influence of specific Coast Guard interventions (involving instrumental, social, or normative factors) can be observed and measured. Over time these observations and measurements can be increasingly fine-tuned for accuracy and effectiveness.

This new digital reality offers the Coast Guard an opportunity to construct a strategy of full-spectrum deterrence that can be measured, continually improved, and credibly communicated. Once the processes outlined above are implemented, investments can be targeted to those aspects of Coast Guard presence that have demonstrated the most consistent benefits or that are needed to engage an especially consequential threat or vulnerability. Command decisions can be complemented by statistically-validated observations and measurements of particular populations' predispositions. A full spectrum of instrumental, social, and/or normative interventions can be deployed to increase the probability of an intended strategic outcome. Progress (or lack of it) can be tracked and interventions adjusted accordingly.

3.12 Engaging Maritime Communities

> *"Departments and agencies that have not historically made wide use of advanced data analytics should make the most out of what the big data revolution means for them and the citizens they serve. They should experiment with pilot projects, develop in-house talent, and potentially expand research and development. From the earliest stages, agencies should build these projects in consultation with their privacy and civil liberties officers."*
>
> *Big Data: Seizing Opportunities, Preserving Values*
> The White House, May 2014

Other than recreational boating and fishing, most of the populations and sub-populations that the US Coast Guard serve seem to have been comparatively slow to adopt the various digital tools involved in social media. This is, however, likely to change in upcoming years. The maritime industry (especially shipping and drilling) already appears to be making serious investments in the Internet-of-Things. There is an urgent need to ensure that the Coast Guard remains fully engaged with these groups as more and more interaction occurs in the digital domain. As outlined heretofore, there are several potential benefits. There are also serious risks.

As a government agency, the Coast Guard is constitutionally bound to protect the personal privacy of US citizens. To generate any of the benefits noted above, the US Coast Guard must preserve and even enhance the sense of trust and mutual respect that characterizes the relationship of the Coast Guard with the communities it serves. Social and normative interventions by the Coast Guard must be understood as facilitating (not manipulating) communities. Social and normative interventions by communities can have a powerful effect to deter bad actors; in the process, the Coast Guard must avoid taking actions that are perceived as threatening personal privacy or diversity of opinion or the free-choice of citizens. As has been

the case in other operational domains, the Coast Guard must be knowledgeable, principled, and self-restrained to engage the digital domain.

The team has concluded that it is crucial for the Coast Guard to begin operating in the digital domain with at least the same strategic commitment as it operates in any of its physical domains, but it should do so in a transparent, exploratory, and incremental fashion. The Coast Guard should organize its approach to the digital domain in such a way that it can learn from tactical failures without threatening sustained strategic engagement. Data security will also be important. Traditional approaches to data access will probably not work. An alternate approach to data security is contained in Appendix C.

The following narrative provides a glimpse into the ability of the ISN framework to gain a greater understanding of the Massachusetts commercial fishermen's SIG's motivations, attitudes and actions regarding preserving the Gulf of Maine fishery ecosystem. The Gulf of Maine fisheries are endangered. The Nature Conservancy has partnered with the Cape Code Commercial Hook Fishermen's Association to improve the long term sustainability of these fisheries.

In another unrelated project involving the SIG, Northeast Fisheries Science Center scientists are unexpectedly partnering with Massachusetts Bay fishing fleets to locate and study the mating habits of cod in order to preserve the species. The scientists understood the mating habits from laboratory experiments but they were unable to find the "spawning haystacks" in the open ocean. Commercial fishing fleets were trying to avoid catching cod. The fishermen told the scientists where to look for the "spawning haystacks." The scientists pinpointed the spawning areas and shared this information to the commercial fishing fleets.

3.13 An ISN Assessment of the Situation

The New England fishermen don't want to catch the Cod. Why? 1) Normative: Catch limit of 1000lbs/year was meant to allow Cod species to recover. 2) Instrumental: Even if caught by accident, cod must be kept; it's unlawful to throw it back. Fined if caught fishing in a no-fishing area. 3) Social: Common ground for Commercial Fishermen and Scientists. Fishermen willingly collaborate with the scientist to help pinpoint where Cod are located, not to fish but to avoid engagement. Based on expert knowledge of fish locations, scientists are able to more accurately place their observing equipment providing the fishermen with very accurate locations to avoid and helping regulators minimize the size of no-fish areas. The strategic motivation is that the fishermen want to improve their profitability and be able to hand down their business to the next generation of commercial fishermen.

Investigating this part of the digital ocean would enable data collections that measure alignment (over time) of the commercial fishing SIGs in this area to these motivations, attitudes and behaviors which would impact: Coast Guard interactions with this group; resource allocations; and relationships between the Coast Guard and the served populations.

3.14 Other Potential Strategic Implications

"Errors using inadequate data are much less than those using no data at all."

Charles Babbage

This analysis began with a focus on deterrence. Partly because of how the original research defined deterrence (more broadly than some other definitions), the benefits and potential pitfalls outlined above clearly have implications for a number of other Coast Guard operational concerns, and especially those related to prevention and performance measurement.

To the extent data can be organized into information and information can be sufficiently contextualized to become knowledge, there is at least the potential to apply enhanced wisdom to a wide array of problems.

To the extent SIG behavior can be measured and to the extent the impact of Coast Guard interventions can be measured against these behaviors, the use of Big Data analytics will presumably be to helpful in better understanding, assessing, and communicating Coast Guard effectiveness in a number of mission areas. But while these further benefits are entirely plausible, they were not the focus of this particular analysis.

The roles of instrumental, social and normative influences are not unique to the Coast Guard. The potential for Big Data analytics in deterrence or for other purposes is not unique to the Coast Guard. The concepts and capabilities outlined above are applicable to several mission sets within the Department of Homeland Security and beyond.

4 RECOMMENDATIONS FOR MOVING FORWARD

The research team suggests that one possible approach for moving forward would be for the Coast Guard to implement a three-to-five year pilot project organized as an independent analysis unit. Fundamental to the pilot project would be the ability to identify and/or recruit and/or develop in cooperation with communities, the civic sector, and the private sector a collaborative capability to gather and analyze "Big Data." For this pilot, we strongly recommend the Coast Guard avoid "ownership" of data-streams or direct access to raw data sources.

As part of the pilot and in cooperation with its civic and private sector partners, the Coast Guard would:

- Develop and distribute practical digital tools to enhance the safety, stewardship, and security of maritime communities in the pilot sector. The Coast Guard and its partners would be explicit regarding how the digital exhausts generated by these tools would be used to inform Coast Guard strategy, engagement, and communications.
- Gather and organize these digital exhausts and combine with other public-facing digital streams to identify and map various SIGs in the pilot sectors, including prospective influence vectors within and between SIGs.
- Develop and deploy early versions of the Data Driven Decision/Deterrence tools previously identified and test utilization of these tools during actual Coast operations.[10]

It is likely this pilot project will require waivers on a range of current Coast Guard policies and regulations that currently complicate use of social media and other sources of Big Data. The project would need to develop forward-learning standards of privacy, security, and other operational requirements.

[10] These tools include: Geo-temporal Visualization Dashboard, Mobile Data Collection Tool, Data Centric Security Framework, and SIG Identification via Social Media.

The pilot project team would need to involve: experienced personnel related to each Coast Guard mission area, communications specialists, software programmers, statisticians, systems engineers, data scientists, product development and tool development professionals, lawyers, and ethicists. It would need to be a public-private team involving personnel from the Coast Guard, other governmental agencies, and both the civic and private sectors. Elements of a pilot project could be organized under Cooperative Research and Development Agreements whereby ownership of many tools, several processes, and most raw data will reside with others in exchange for guaranteed Coast Guard access to defined and "scrubbed" data.

The team recommends intelligence personnel not be involved in the pilot project. There are undoubtedly intelligence implications of Data Driven Deterrence. But intelligence purposes, at least as typically understood today, should not shape the design or conduct of the pilot. To do so would threaten to undermine the trust and mutual respect on which full-spectrum deterrence is perceived to depend.

The pilot project would also need to involve significant training and education related to use of the Data Driven Deterrence tools, appropriate interpretation of information for decision-support, and assurance of legal and ethical behavior.

The success or failure of the pilot would depend a great deal on how the following questions are answered:

1) Can the U.S. Coast Guard and its partners develop and deploy digital tools that have sufficient utility to maritime communities that individuals will share their digital exhausts with a full understanding of how the data will be used to examine and engage populations and sub-populations?
2) Can the civic and private sector partners gather sufficient data to provide the US Coast Guard with an accurate understanding of behavior by and normative values of Social Identity Groups from which USCG can derive meaningful influence vectors to inform operational decisions?
3) Does the piloted practice of Data Driven Deterrence demonstrate any improved ability to measure the interface between US Coast Guard interventions and actual sector outcomes and, even more ambitiously for a three-to-five year study, is it possible to demonstrate that the early practice of Data Driven Deterrence actually improves sector outcomes?

Additional details of the proposed pilot project are contained in Appendix A.

5 BROADER STRATEGIC IMPLICATIONS/CONSIDERATIONS

"Errors using inadequate data are much less than those using no data at all."

Charles Babbage

This analysis began with a focus on deterrence. Partly because of how the original research defined deterrence (more broadly than some other definitions), the benefits and potential pitfalls outlined above clearly have implications for a number of other Coast Guard operational concerns, and especially those related to prevention and performance measurement.

To the extent data can be organized into information and information can be sufficiently contextualized to become knowledge, there is at least the potential to apply enhanced wisdom to a wide array of problems.

To the extent SIG behavior can be measured and to the extent the impact of Coast Guard interventions can be measured against these behaviors, the use of Big Data analytics will presumably be helpful in better understanding, assessing, and communicating Coast Guard effectiveness in a number of mission areas. But while these further benefits are entirely plausible, they were not the focus of this particular analysis.

The digital domain outlined above is in many ways analogous to the rather sudden emergence of a new ocean. The opportunities and challenges presented are each enormous.

6 ROADMAP

Building Big Data capabilities demands a change of mindset from organizations, from the idea that it is just about technology acquisition to the understanding that big data projects require organizational change.

The roadmap in Figure 12 describes one typical path for an organization to adopt Big Data capabilities within its decision-making models. Before the CG can adopt a roadmap, additional analysis is required to determine affordability and feasibility, including:

- Evaluating data storage requirements and costs associated with upgrading IT infrastructure for big data collection, aggregation, collection, storage, analysis and reporting.
- Estimating costs for acquiring personnel with specialized skill sets such as data analysts, data engineers, data architects and scientists and consultants
- Estimating costs for implementing data security controls to protect big data.

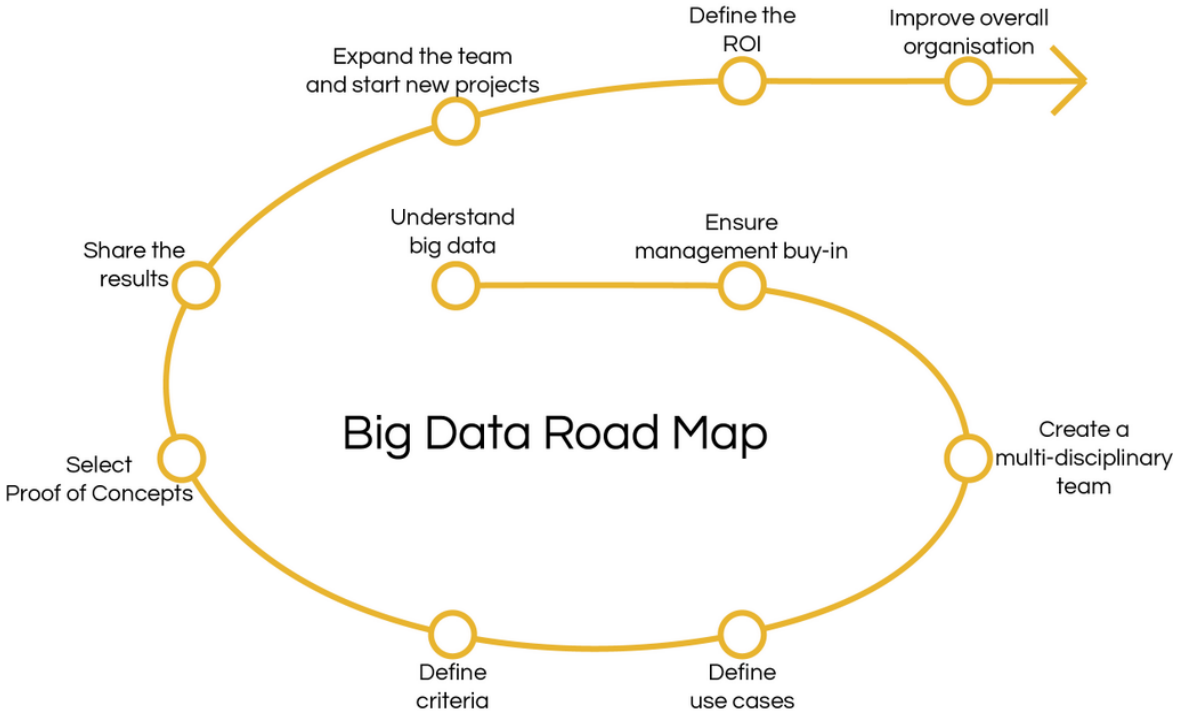


Figure 12. Big Data Road Map designed by datafloq.com.

(This page intentionally left blank.)

APPENDIX A. DETAILS OF PROPOSED PILOT PROJECT

Conceptually, the potential pilot project would consist of assets and resources distributed across the Coast Guard, specifically organized and authorized to conduct activities to enhance Coast Guard mission-achievement in the Digital Ocean. Lessons learned – both positive and negative – will be utilized to inform policy, strategy, tactics, techniques, and procedures across the Coast Guard.

A.1 Value Proposition

At the center of the business canvas is the value proposition of Data Driven Engagement through the pilot project. It contains how the pilot project would create value for its Coast Guard customers:

- It would provide and constantly update the tools required for the Coast Guard to have a multidimensional Data Driven Engagement capability.
- It would build institutional capacities for the Coast Guard to implement Big Data solutions that are ethical and effective.
- It would delineate intrapreneurship mechanisms for the Coast Guard to be able to follow the exponential growth of Data Driven technology.
- It would manage an app store approach to behavioral intervention and deterrence.

A.2 Key Activities

The pilot project could provide this value by performing the following key activities:

- Big Data Analysis.
- Data Capture for Instrumental, Normative, and Social dimensions of deterrence/influence.
- Tech development for the whole Data-to-Deterrence cycle: Sensors, data capture, analysis and modeling, visualization and sense-making, command and control, and evaluation.
- Technology evaluation and procurement capabilities.
- Exploitation teams (keep successful projects running).
- Intrapreneurship mechanisms to encourage exploration of new methods to follow the exponential pace of change.

A.3 Key Partners

To create value, the pilot project would need to effectively interact with the following key partners:

- DHS S&T.
- The private tech sector (Google, Facebook, etc. but also small startups working on these problems).
- NPS/CHDS.
- Other Big Data organizations in government.
- Amazon cloud or other cloud computing facilities.

A.4 Key Resources

Most resources needed are translatable into labor with the exception being data storage and procurement of new technology. The following resources will be required:

- Labor Technological: Application development capability.
- Labor Quantitative Science: An in-residence quantitative science capability to apply the most novel methods and models to Deterrence and Decision Making practices.
- Labor Design: Sensemaking designers to produce high-quality interfaces and user-friendly products.
- Labor Tech Evaluation: A team of tech evaluators capable of rapidly following technology trends and evaluate potential applications for Coast Guard missions.
- Labor Innovation: Intrapreneurship teams exploring new realms following lean startup practices.
- Mobile/sensor technology and displays.
- Data Farms or cloud computing.

A.5 Cost Structure

Because most of the resources needed are labor related, most of the costs are also associated with labor:

- Labor for tech development, quantitative science, design for Human-Computer Interaction, for tech evaluation, for innovation teams.
- Data farming and data hosting costs.
- Mobile/sensor technology procurement (e.g., wearables like the Apple Watch).

A.6 Customer Segments

The pilot project creates value for two kinds of customers: internal and external.

Internal

The internal customers for the pilot project are key decision makers both in the Coast Guard Headquarters and DHS who have strategic needs to improve their Data-to-Decision cycles.

Also, operational decision-makers in Coast Guard sectors and regions and DHS components will benefit from the capabilities provided by the pilot project.

External

Outside of the Coast Guard, the pilot project would create value to Social Identity Groups that constantly interact with the Coast Guard by creating new channels of interaction (e.g., fishermen, boating community, cruise ship tourists, coastal population, etc.)

A.7 Channels

Through the pilot project, the Coast Guard would interact with internal and external customers through three channels:

- Face-to-face briefings.
- Technology dashboards and visualizations.
- Mobile "apps."

A.8 Customer Relationships

Customers, interacting through those channels, would establish the following relationship with the pilot project:

Community curation

The pilot project would create and manage platforms that foster community interaction with Social Identity Groups. For example, this is exactly Facebook's core relationship with its users.

Data analysis

Data analytics are the key reason why data-driven organizations create platforms. By capturing the digital pheromones produced as a byproduct of technology-mediated human interactions, it is possible to gain insight about changes in the shape and form of cultural and normative attitudes, beliefs and behaviors of those communities. This is what advertisers look for when they pay millions of dollars to social media companies to communicate with those social identity groups. The behavioral intervention, in those cases, is oriented at producing sales.

Visualization provider

As we capture more digital pheromones, it becomes possible to build displays that allow the visualization of social changes. A good metaphor to explain the value of social physics dashboards is how when Air Traffic Control communicates through a technology (the radio) with pilots and provides instructions such that the controller can perceive the impact of this communication using a radar display screen, i.e. either the pilot altered his/her heading or did not. We can know that (and aggregate and analyze it the way the FAA does) because this human interaction is technologically mediated.

As more and more interactions take place online, social interactions also become technologically mediated and thus susceptible to the same kind of visualization. We can visualize conversations online around a particular event or geotemporal information to cite two current capabilities developed by this project for the pilot project.

Decision-making support and tracking

Ginni Rometty, CEO of IBM recently stated that "In the future, every decision that mankind makes is going to be informed by a cognitive system like Watson," she said, "and our lives will be better for it." The pilot project will assist in the construction of cognitive and decision-making systems for the Coast Guard.

A.9 "Revenue" Streams

The positive impact a pilot project and its Data Driven Engagement model would have on Coast Guard capabilities is reflected on the "revenue stream" box of the canvas. In this case, revenue is not understood as dollars like in the case of a for-profit organization, but in the form of the benefits the Coast Guard would obtain from the operation of the pilot project.

There are at least four lines of "revenue" generated by the pilot project:

1) Improved CG mission effectiveness and efficiencies.

As the Coast Guard creates or procures the necessary tools to capture, model, analyze, visualize and decide about digital pheromones created by technologically mediated social interaction, it will develop experimental capabilities. These experimental capabilities allow for experimental trials to demonstrate if the underlying hypotheses of programs and policies are true or not. What this means is that if a specific program claims that by doing X the effect will be Y. Today, it is often the case that it is impossible to validate if that claim was true or not.

An odd, but great example to showcase the way technology-mediated interactions allow for this analysis of effectiveness and efficiencies in ways not possible without data is Google's famous 41 shades of blue. In the words of Marissa Mayer, former VP of Product Search:

> *"Turns out Google was using two different colors of blue, one on the homepage, one on the Gmail page. To find out which was more effective so they could standardize it across the system, they tested an imperceptible range of blues between the two. The winning color, according to dozens of charts and graphs, was not too green, not too red. "It's interesting to see how you can change the way that people respond to the Web in ways that are not intuitive."*

2) Improved maritime awareness.

Given the amount of data available and the cognitive limitations of the human brain, pattern recognition in the era of exponential data is a difficult task. The pilot project will provide "nowcasting" capabilities for the Coast Guard by creating cognitive systems that help decision makers to make better choices, with improved situational awareness informed by near real time data collections.

For example, today, the Coast Guard performs multidimensional deterrence through multiple channels, but it is often not aware of it, nor does it have any methods to observe the impact of those actions. The pilot project would change that by giving to the Coast Guard the same set of tools that big corporations already have, and that criminal organizations are learning to exploit.

3) A better return on multidimensional intervention/influence.

Many problems that the USCG confronts arise from questions that today, through an experimental approach, can be answered. Would action X increase or decrease the amount of boating accidents? With the right tools for data capture, analysis, and visualization, there is a way of answering that question in near real time.

4) Better relationships with key affected populations in the implementation of actions to fulfill the Coast Guard's 11statutory missions.

The pilot project would multiply the channels through which the Coast Guard would interact with the Social Identity Groups that are concerned by its actions. Not only would the amount of interactions increase, but also interaction quality, as the Coast Guard learns from the common experience of these groups in a systemic way. In a way, it would provide a vehicle for the Coast Guard to democratize and crowdsource its learning capacities.

APPENDIX B. SOCIAL MEDIA SIG IDENTIFICATION & MONITORING

This analytics methodology identifies and monitors social identity groups (SIGs) through analysis of unstructured information sources such as social media posts, web blogs, news article content, email content, etc.

B.1 Need/Significance

Identifying and monitoring SIGs is a foundational component of establishing effective data-driven deterrence. A deterrence vector of instrumental, social, and normative interventions is likely to impact various SIGs in different ways. Therefore, monitoring the effect of all three of these vectors on performance – as well as improving future deterrence vector composition – requires knowledge not only of the relevant mission-related SIGs, but also of who composes those SIGs and how those SIGs change over time.

B.2 Background

Social media monitoring (SMM) tools have matured to perform unstructured information analysis on more than just social media. Many now monitor web blogs, news articles, and other rapidly-growing text-based public information sources. SMM tools are built to monitor a specific set of words to examine trends and sentiment related to those words. This approach is valuable to private sector companies as they monitor static keywords (e.g., Wal-Mart, Dell, iPhone 5S) and manage their company's brand or product reputation.

However, maintaining situational awareness in USCG mission areas such as drug interdiction, homeland security, migrant interdiction, and living marine resources requires monitoring a shifting set of actors that quickly alter their behavior and language to avoid detection. In such contexts, existing SMM tools are inadequate since they are unable to effectively track both the shifting member population of SIGs and the shifting language used by SIG members.

For USCG data driven deterrence, a new unstructured information analysis is needed to identify and monitor the SIGs that impact various USCG mission areas.

B.3 Approach

Both language and topics of conversation are collectively emergent. That is, they require *communities* to have useful meaning. Yet communities, as well as the language and conversations that sustain them, are neither simple nor static constructs.

B.3.1 Multi-Scale and Overlapping Communities

Any individual is a member of many communities that impact their language and conversation. For example, Houstonians say "washateria" (i.e., laundry mat), Texans say "y'all" (i.e., you all), southerners say "fireflies" (i.e., lightning bugs), and Americans say "dude" (i.e., bloke, mate). These geographical communities exhibit a clear multi-scale pattern (i.e., a city exists within a state, a state within a region, etc.). However, a Houstonian may also ascribe to a faith tradition, maintain industry association memberships, or share similar demographic characteristics with communities that influence their language and conversation

interests. Each of these communities is likely to have other members residing in different cities, states, or even nations – that is, they can overlap with geographical communities at many scales. Therefore, communities can be multi-scale and can overlap in ways that do not respect other communities' scales as depicted in Figure B- 1.

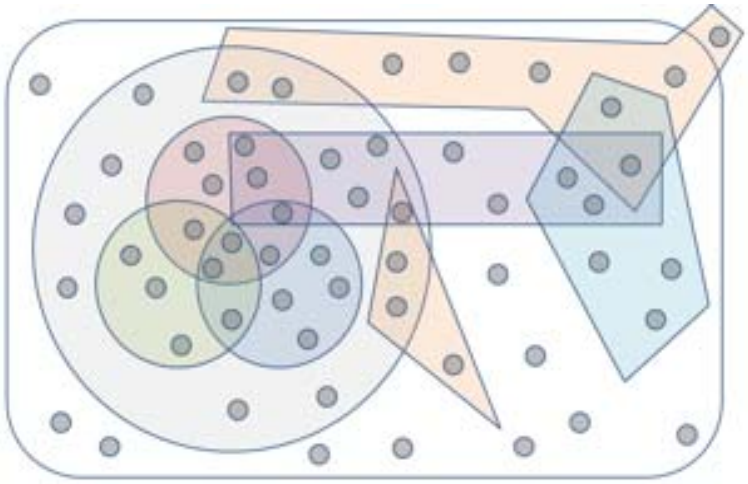


Figure B- 1. Multi-scale and overlapping Social Interest Groups (SIGs).

B.3.2 Community Dynamics

In addition to being multi-scale and overlapping, the scales and boundaries for many communities are dynamic. They can form, grow, nest, shrink, merge, split, and disappear. For example, DHS is concerned with the evolving threat from terrorist organizations. As some terrorist groups come and go or their members move, these communities dynamically overlap with each other – as well as with geographical communities from local to international scales across the globe.

B.3.3 SIG Identification

As users generate more posts, their content can be used to assess their degree of alignment with others' posts. Users that post the same groups of SIPs and discuss similar events/issues can be clustered into SIGs. Over time, their activity can give rise to dynamic SIG alignment.

Figure B- 2 and Figure B- 3 depict how monitoring the set of SIPs can be used to determine a set of SIGs that users share with others. The figure also shows how relationships between SIGs (supersets, subsets, or intersecting sets) can be monitored with this approach. For example, local, regional, and national-level political communities may use different language to discuss similar issues. Similarly, when words such as wind, power, flood, and damage, are posted by several users not previously aligned with disaster-related SIGs as well as by others that have been posting words like plan, mitigate, response, and recovery for many years, it may signal that a longstanding disaster relief worker SIG is now intersecting with a new disaster-affected SIG.

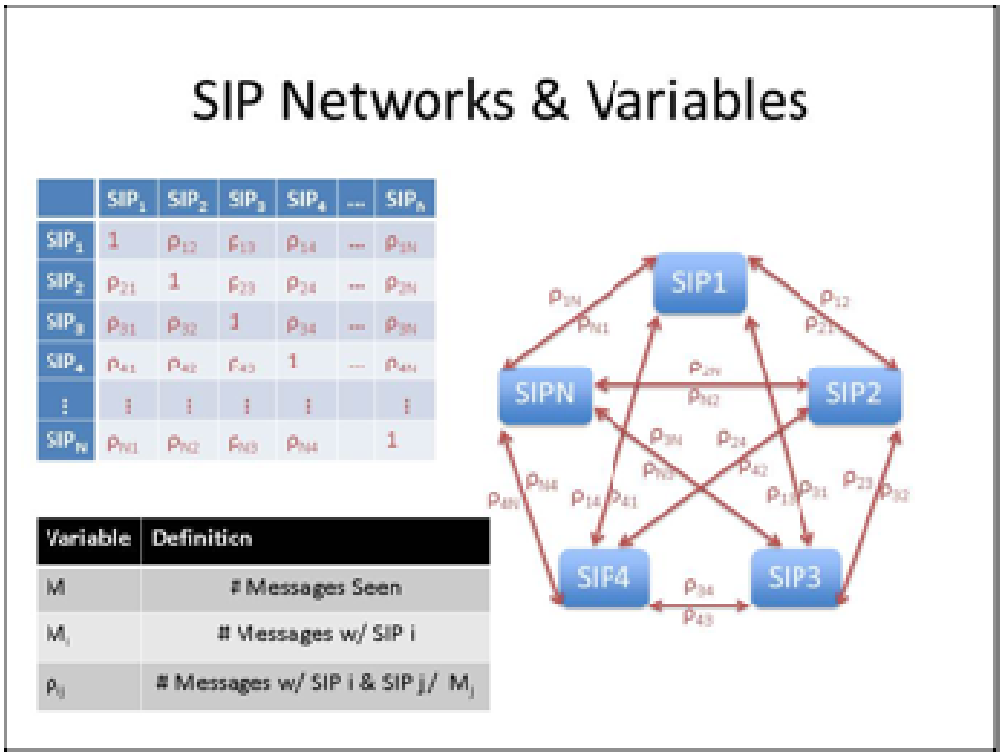


Figure B- 2. Generic SIP network.

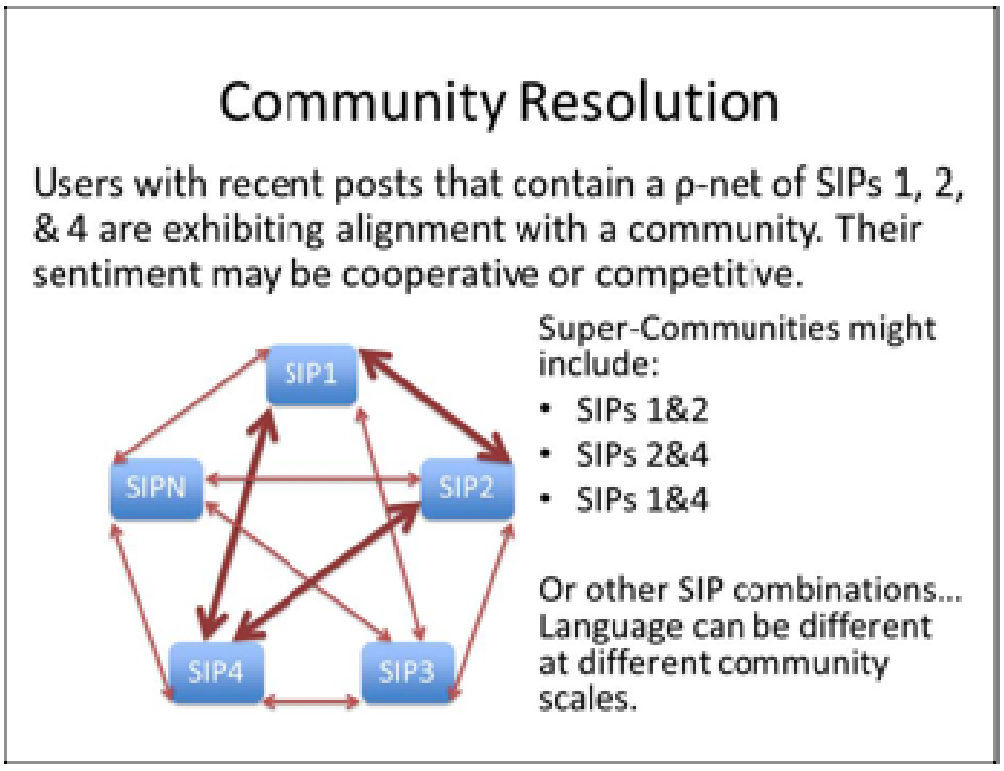


Figure B- 3. SIG identification.

While SIP use will be used as the primary SIG identification mechanism, several statistical network measures will also be used to examine relationship connectivity, proximity, clustering, redundancy, and other measures. These measures will be applied to inform SIG alignment based on various explicit interactions (e.g., friends, follows, re-shares, comments, likes).

B.3.4 Word Frequencies

SIPs can be used to determine a set of communities that users share with others. Relationships between these communities (supersets, subsets, or intersecting sets) can also be monitored with this approach. For example, local, regional, and national-level political communities may use different language to discuss similar issues. Similarly, when words such as wind, power, flood, and damage, are posted by several users not previously aligned with disaster communities as well as by others that have been posting words like plan, mitigate, response, and recovery for many years, it may signal that a longstanding disaster relief worker community is now intersecting with a new disaster-affected community.

Word frequencies can be used to determine not only which words are SIPs, but also how impactful of a SIP those words are for both individuals and communities. Figure B- 4 is a power law graph showing the reverse rank-ordered use of English (American) language words. Figure B- 5, shows the frequency of words used within a set of fishing community interviews superimposed on this power law distribution.

Word	Occurs	Freq%
the	22038615	4.89747%
be	12545825	2.78796%
and	10741073	2.38691%
of	10343885	2.29864%
a	10144200	2.25427%
in	6996437	1.55476%
to	6332195	1.40715%
have	4303955	0.95643%
to	3856916	0.85709%
it	3872477	0.86055%
I	3978265	0.88406%
that	3430996	0.76244%
for	3281454	0.72921%
you	3081151	0.68470%
he	2909254	0.64650%
with	2683014	0.59623%
on	2485306	0.55229%
do	2573587	0.57191%
say	1915138	0.42559%
this	1885366	0.41897%
they	1865580	0.41457%
at	1767638	0.39281%
but	1776767	0.39484%
we	1820935	0.40465%
his	1801708	0.40038%
from	1635914	0.36354%
that	1712406	0.38053%
not	1638830	0.36418%
n't	1619007	0.35978%
by	1490548	0.33123%
she	1484869	0.32997%
or	1379320	0.30652%
as	1296879	0.28820%
what	1181023	0.26245%
go	1151045	0.25579%
their	1083029	0.24067%
can	1022775	0.22728%
who	1018283	0.22629%
get	992596	0.22058%
if	933542	0.20745%
would	925515	0.20567%
her	969591	0.21546%
all	892102	0.19824%
my	919821	0.20440%
make	857168	0.19048%
about	874406	0.19431%
know	892535	0.19834%

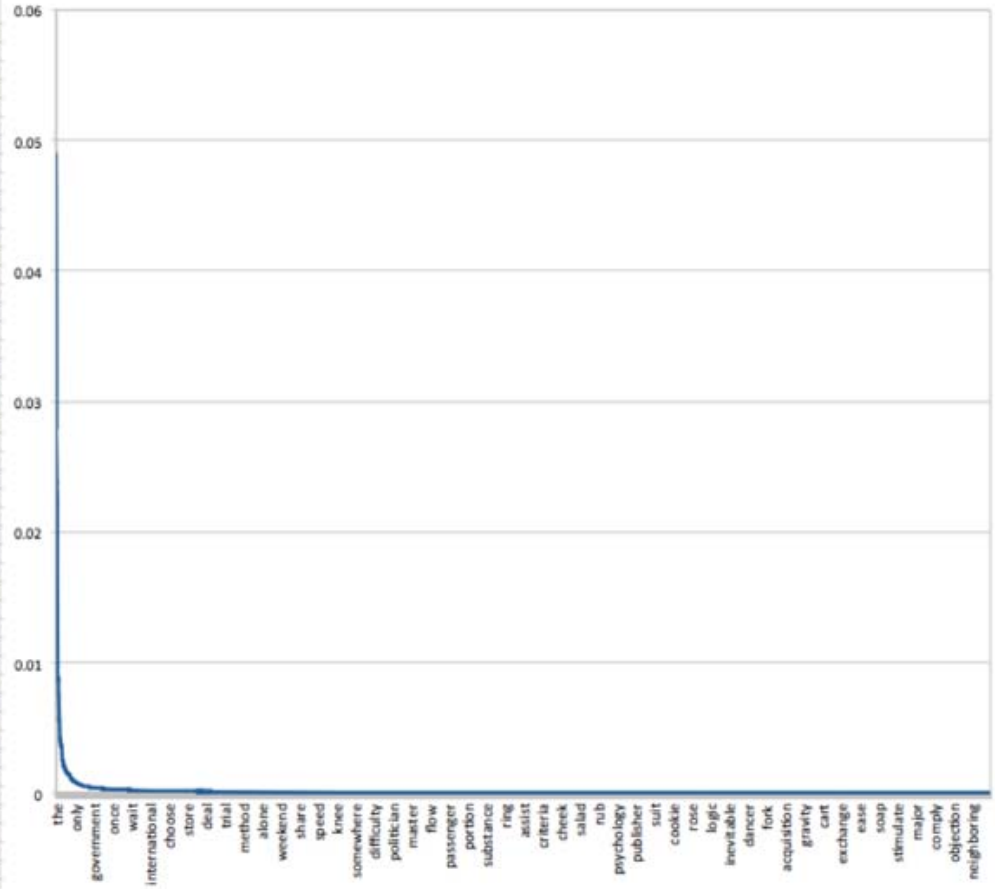


Figure B- 4. Reverse rand-ordered power law graph of the English language words.

There are several spikes for words that are somewhat unique to this community. Examining the ratio of group word frequency to general American English word frequency, we can quickly identify which SIPs are most abnormal. At the top of the list are words including “fishing,” “boat,” “fish,” and “catch.” These correspond to spikes in the word frequency graph, which visually depict how certain words can point to the likelihood of certain terms or topics being relevant within this community.

	occurs	freq_group	freq_glob	ratio
fishing	313	0.005095	4E-05	126.6395
boat	206	0.003353	7.1E-05	47.03508
fish	263	0.004281	9.2E-05	46.43107
yeah	475	0.007731	0.00023	33.65075
catch	143	0.002328	0.00015	15.35458
mean	325	0.00529	0.00054	9.828523
right	241	0.003923	0.00046	8.600215
lot	180	0.00293	0.00038	7.774974
well	324	0.005274	0.00092	5.763143
so	561	0.009131	0.00168	5.431261
know	636	0.010352	0.00198	5.21924
there	508	0.008268	0.00174	4.742755
where	137	0.00223	0.00047	4.694638
no	252	0.004102	0.00089	4.588917
like	347	0.005648	0.00126	4.467938
you	1802	0.02933	0.00685	4.283682
much	134	0.002181	0.00054	4.014107
then	294	0.004785	0.00121	3.958606
here	209	0.003402	0.00092	3.712723
out	330	0.005371	0.00151	3.561833
now	291	0.004736	0.00135	3.517207
i	1855	0.030193	0.00884	3.415275
because	204	0.00332	0.00097	3.407198
them	312	0.005078	0.00151	3.371193
down	151	0.002458	0.00073	3.357507
just	309	0.005029	0.00151	3.339561
back	151	0.002458	0.00082	3.00669
what	468	0.007617	0.00262	2.902438
good	139	0.002262	0.00079	2.876207
they	727	0.011833	0.00415	2.854276
think	296	0.004818	0.00172	2.805481
up	293	0.004769	0.00177	2.697642
we	660	0.010743	0.00405	2.654758
when	243	0.003955	0.00151	2.622716
how	184	0.002995	0.0012	2.500868
that	1160	0.018881	0.00762	2.476357
all	296	0.004818	0.00198	2.430259
more	168	0.002734	0.00115	2.377629
your	192	0.003125	0.00147	2.131972
or	397	0.006462	0.00307	2.108147
would	253	0.004118	0.00206	2.002223

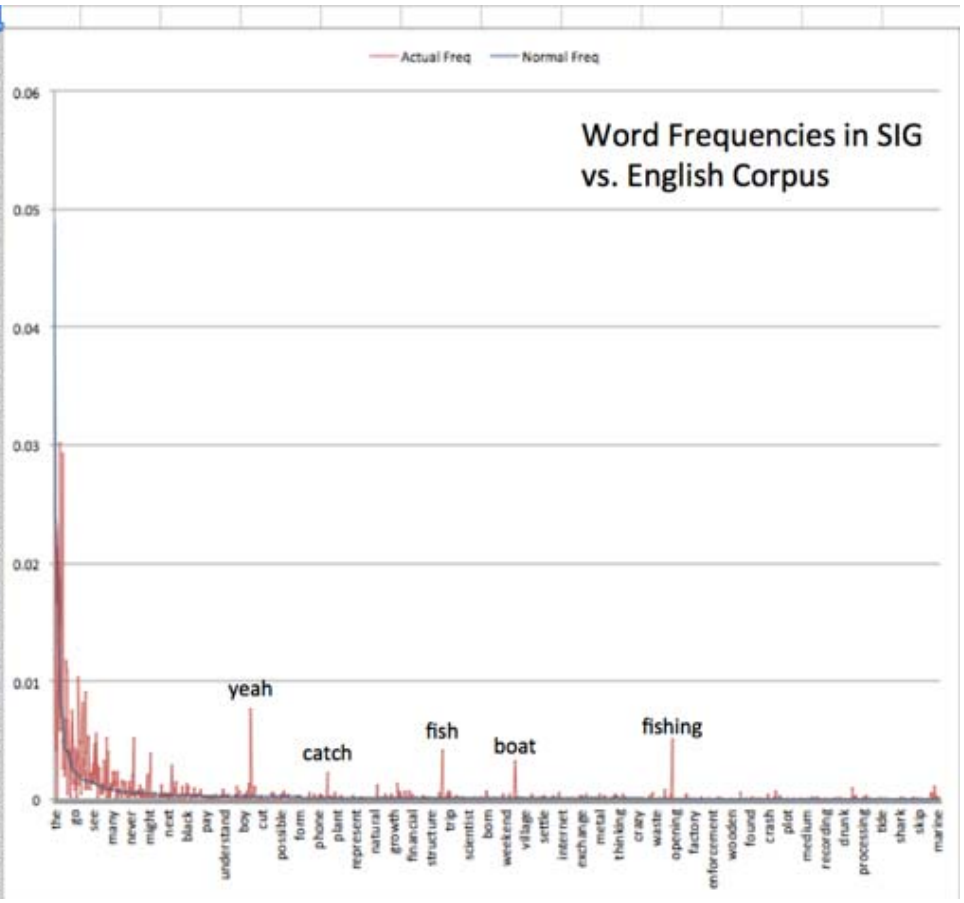


Figure B- 5. Fishing community word frequencies overlaid on the normalized English word count.

Emergent communities can be identified by examining individuals' use of words and clustering users together based on SIP usage.

APPENDIX C. DATA-CENTRIC SECURITY

The approach taken by Data-Centric Security (DCS) represents a fundamental shift. In the past, information security managers have largely attempted to obtain security by controlling the evolving mixture of infrastructures and systems. DCS provides security down to the data object level so that, no matter what networks, devices, applications, or people are involved, you retain full control over access to your information. It uses a proven set of scalable, data-centric information security management processes that bridge the seams among infrastructures and applications from yesterday, today, and tomorrow.

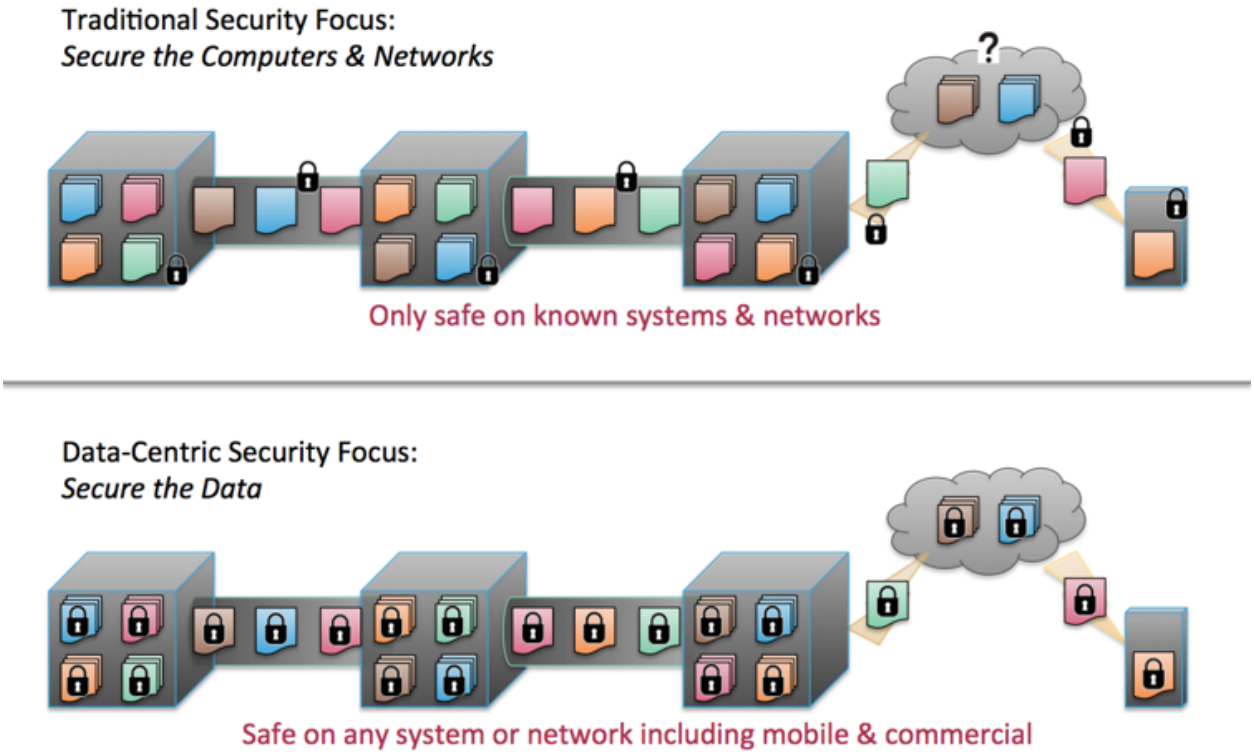


Figure C- 1. The advantage of a data centric security approach.

C.1 DCS Technology

To create a DCS data object (a file), you first authenticate on a trusted DCS server (usually hosted in your own organization) to obtain keys for securing your data. A client DCS library uses RSA & AES encryption as well as other security measures to transform your data into a data object that can only be accessed with a specific set of keys. At any time, you may define – or redefine – the access control list, permissions, and business rules for your data. The DCS data object remains on your computer – the keys, access control list, and permissions are maintained and managed remotely on the trusted DCS server (which never sees your data). From then on, no matter where your DCS data object is or where it goes, your data is safe. Recipients may save, copy, and distribute the DCS data object via any means – whether it is transferred via flash drives, copied behind enterprise firewalls, emailed, or even stored in a public cloud for mobile device access. To access the data within the DCS data object, other users must authenticate on the trusted DCS server and have appropriate permissions as well as meet any business rules you define.

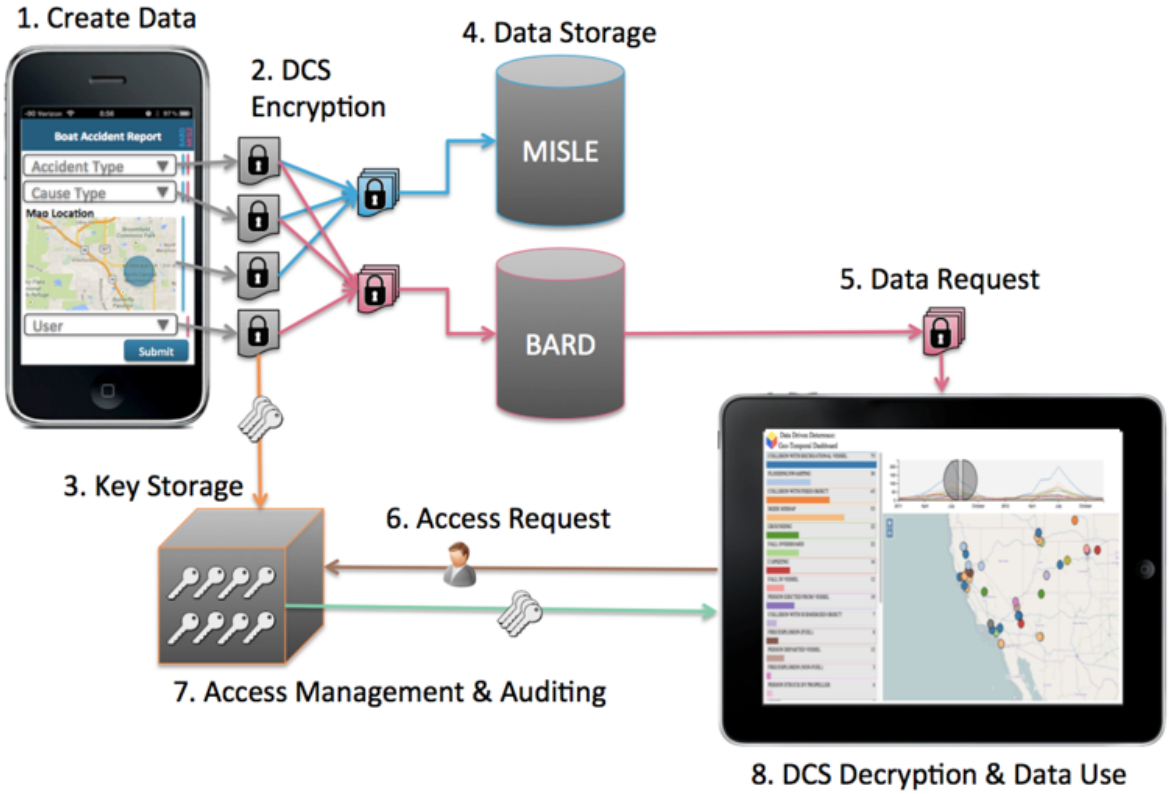


Figure C- 2. Protection starts at data generation and continues to storage and retrieval.

C.2 Supporting Existing Infrastructures & New Capabilities

The DCS platform can integrate with existing infrastructures and applications. DCS servers can be configured to use multiple identity management providers, leveraging existing systems for single sign-on, security cards (e.g., CAC/PIV), and biometrics. Regardless of the level of security needed, the DCS platform can ensure that the appropriate access requirements are met.

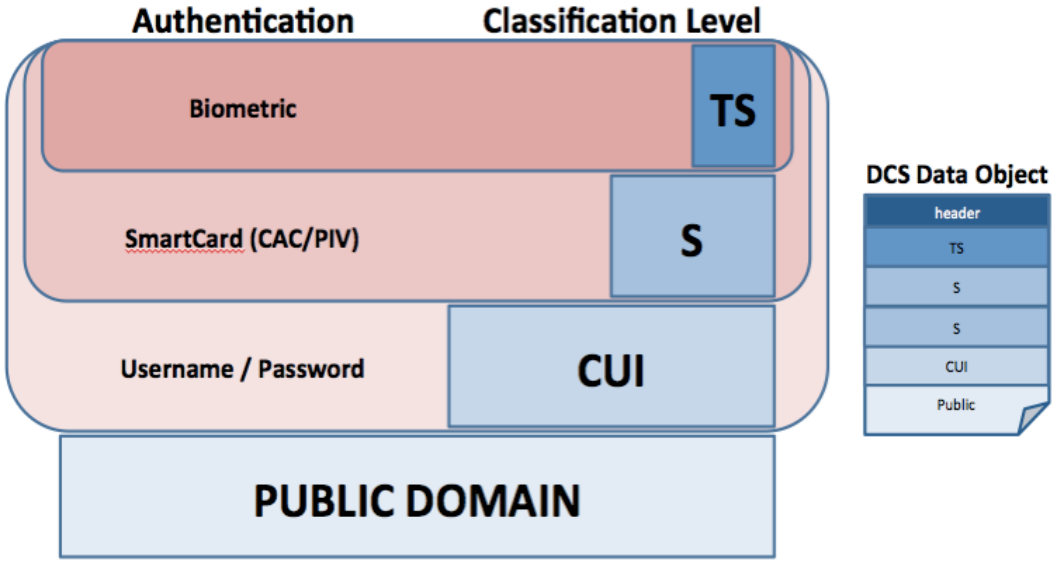


Figure C- 3. DCS configuration.

The DCS platform ensures that everyone has access to the information they need to know and *only* the information they need to know. Multiple levels of protection can co-exist securely within the same DCS data object. Furthermore, creators can continue to manage access permissions remotely so as needs change, access to information can be granted, modified, or even revoked. That is, even after files have been distributed, access and protection levels can be altered and those who have already viewed the information can be notified of the change. Such data-centric information security management processes make several new capabilities possible, including:

- Post-distribution withdrawal of access to old versions.
- Digital tear-sheets within your documents.
- Escalation & de-escalation of data classification.

C.3 Applying Additional Business Rules

The DCS platform allows additional business rules that provide unprecedented control over your data. Beyond just using access control lists and permissions to specify who can access your data, creators can also define infrastructure requirements, geo-fencing, and time windows. For example, in addition to identity management requirements, creators could specify that a DCS data object can only be accessed via iOS mobile devices in the Washington, D.C. metro area between the hours of 8am and 5pm. Even recipients on the access control list and with appropriate permissions would not be able to access the data when on the wrong device, outside the permitted area, or after working hours.

Figure C- 4. Geo-fencing and time windows.

C.4 Security Auditing

The Secure Objects platform provides security auditing along with data health monitoring and customizable automated alerts. Creators and managers can set up alerts to be notified whenever data is accessed, when unauthorized access attempts are made, and/or when the integrity of data has been compromised. Policy controls and audit groups enable DCS servers to create detailed audit trails of every attempt to view your data. An administrator console provides reports to security administrators with information such as number of unopened DCS data objects, frequency of attempted unauthorized access, number of DCS data objects created, and much more. Administrators are able to track the DCS activity of each individual user, making Secure Objects creators accountable for all of the information they transmit within and outside the organization.

C.5 Standards

DCS uses a FIPS-140-2 validated AES256 encryption library, affording the use of the DCS platform within the United States government—across both military and civilian agencies.

C.6 The User Experience

The DCS platform is not a specific application with a user interface; rather, it is an application programming interface (API) that provides access to its collection of scalable, customizable, data-centric information security management processes. This API can provide a simplified user experience by integrating into any application.

For your applications, the DCS platform API can be used in ways that range from completely transparent to highly integrated. Many organizations may opt to use add-ins or security wrappers to ensure data remains protected without the need to modify existing software code. This makes integrating your existing systems easy and cost-efficient while taking the pain out of any future upgrades or maintenance. Other organizations may desire to customize how their applications use the API for high-security environments. Regardless, the API allows a scalable, data-centric set of information security management processes to be quickly and seamlessly integrated into organizations' new and existing applications.

C.7 The Result: Information Security in a Complex & Growing Network

A data-centric information security management platform provides a way to address the complex information sharing challenges among a growing network of organizational partners. Instead of securing the exponentially increasing number of interactions among various infrastructures and applications, we can focus on ensuring that the data that must move among them remains secure and safely controlled by its creators and/or managers.

This approach provides more for less; the cost of information security will be drastically reduced – in terms of both data management and user experience. As data-centric processes are more widely implemented, we will not only see greater consistency, automation, and monitoring efficiency, but we will also see changes to our infrastructures and applications that reduce the information security management burden. Furthermore, users will be freed from the need to keep track of redundant networks and applications. As security evolves to be more data-centric, personal mobile devices can finally be used to do the jobs necessary for meeting sensitive organization needs.

Despite the reality of increasing change and complexity, ultimately, a data-centric information security platform provides what we are seeking. We remain in control of our data – always and everywhere. No matter what infrastructures or applications are being used and no matter where our data travels or is stored, we can be assured that it is appropriately secured.

(This page intentionally left blank.)

APPENDIX D. LIST OF REFERENCES

1. United States Coast Guard Cyber Strategy, United States Coast Guard, June 2015

2. Deterrence and the United States Coast Guard: Enhancing Current Practice with Performance Measures, Research and Development Center, March 2012

3. Phillips, Donald and Loy, James M., Character in Action: The US Coast Guard on Leadership, November 2015.

4. Ostrom, Elinor and Gardner, Roy and Walker, James, Rules, Games and Common Pool Resources, University of Michigan Press, 1994

5. Ostrom, Elinor, Proceedings of the National Academy of Sciences, Volume 4, Number 39, September 25, 2007

6. Ostrom, Elinor, Beyond Markets and State: Polycentric Governance of Complex Economic Systems (Nobel Prize Lecture), December 2009

(This page intentionally left blank.)